SSIFIED

556

ATION CENTER

ICAL INFORMA

ME NDRIA, VIRGINIA

CLAS D

WEAPON SYSTEM EFFECTIVENESS
[illegible]USTRY ADVISORY COMMITTEE [illegible]

FINAL REPORT OF
TASK GROUP IV

COST-EFFECTIVENESS OPTIM[illegible]

(TECHNICAL SUPPLEMENT)

[illegible]

HEADQUARTERS [illegible]
ANDREWS AIR FORCE BASE [illegible]

WEAPON SYSTEM EFFECTIVENESS

INDUSTRY ADVISORY COMMITTEE (WSEIAC)

FINAL REPORT

of

TASK GROUP IV

COST-EFFECTIVENESS OPTIMIZATION

(TECHNICAL SUPPLEMENT)

FOREWORD

This is Volume III of the final report of Task Group IV of the Weapon System Effectiveness Industry Advisory Committee (WSEIAC). It is submitted to the Commander, AFSC in partial fulfillment of Task Group IV objectives cited in the committee Charter. The final report is contained in three separate volumes:

> Volume I presents a summary of the principles of cost-effectiveness analysis, conclusions and recommendations.
>
> Volume II contains a discussion of the specific tasks required to conduct a cost-effectiveness analysis, emphasizing procedural and analytical techniques.
>
> Volume III consists of a technical supplement illustrating some of the methodology appropriate to cost-effectiveness analysis.

The membership of Task Group IV was as follows:

Mr. H. S. Balaban	ARINC Research Corporation
Mr. C. Davenport	Rocketdyne - Division of North American Aviation, Inc.
Mr. B. W. Davis	Battelle Memorial Institute
Dr. A. S. Goldman	GE TEMPO
Dr. G. R. Herd (Technical Director)	Booz-Allen Applied Research, Inc. (Now with Kaman Aircraft Corporation)
Mr. T. K. Jones	The Boeing Company
Mr. M. Kamins	The RAND Corporation
Mr. R. E. Kuehn	IBM - Space Guidance Center
Dr. G. W. Morgenthaler	The Martin Company
Maj Frank H. Moxley, Jr. (Government Chairman)	Headquarters, AF Systems Command
Mr. J. J. Naresky	Rome Air Development Center

Other task group reports submitted in fulfillment of the committee's objectives are

AFSC-TR-65-1	Final Report of Task Group I "Requirements Methodology"
AFSC-TR-65-2	Final Report of Task Group II "Prediction - Measurement"

AFSC-TR-65-3	Final Report of Task Group III "Data Collection and Management Reports"
AFSC-TR-65-5	Final Report of Task Group V "Management Systems"
AFSC-TR-65-6	Final Summary Report "Chairman's Final Report"

Publication of this report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.

APPROVED

William F. Stevens

William F. Stevens, Colonel, USAF
Chief, Systems Effectiveness Division
Directorate of Systems Policy
DCS Systems

WSEIAC CHARTER

In order that this report of Task Group IV may be studied in context with the entire committee effort, the purpose and task group objectives as stated in the WSEIAC Charter are listed below

Purpose

The purpose of the Weapon System Effectiveness Industry Advisory Committee is to provide technical guidance and assistance to AFSC in the development of a technique to apprise management of current and predicted weapon system effectiveness at all phases of weapon system life.

Task Group Objectives

Task Group I - Review present procedures being used to establish system effectiveness requirements and recommend a method for arriving at requirements that are mission responsive.

Task Group II - Review existing documents and recommend uniform methods and procedures to be applied in predicting and measuring systems effectiveness during all phases of a weapon system program.

Task Group III - Review format and engineering data content of existing system effectiveness reports and recommend uniform procedures for periodically reporting weapon system status to assist all levels of management in arriving at program decisions.

Task Group IV - Develop a basic set of instructions and procedures for conducting an analysis for system optimization considering effectiveness, time schedules, and funding.

Task Group V - Review current policies and procedures of other Air Force commands and develop a framework for standardizing management visibility procedures throughout all Air Force commands.

ABSTRACT

A discussion of optimization which amplifies the material in Volume II, Section IV is presented. Optimization principles, criteria and checklists, as well as a summary of various applicable techniques is included. A series of six examples are described covering a number of critical aspects of cost-effectiveness analysis in considerable detail. Treated in the examples are: (1) Optimization of effectiveness based on reliability, maintainability, performance, and cost; (2) Allocation of reliability requirements among subsystems; (3) Payload allocation among three subsystems based on a fixed weight constraint; (4) Determination of best checkout routine for a limited pre-launch test; (5) Optimization of availability for a complex system; and (6) Trade-off study between site hardening and dispersal for a missile system.

TABLE OF CONTENTS

ILLUSTRATIONS

TABLES

SECTION I

INTRODUCTION

Volume II of the final report of Task Group IV discussed cost-effectiveness analysis in general terms. In this volume we shall elaborate on some of the principles set forth earlier. We shall do this by means of relatively detailed examples which will illustrate some of the uses of models during the Definition and Acquisition Phases of a system life.

As noted earlier, the principle use of a model during the Definition Phase is as an aid in selecting between alternative (competing) system configurations, all of which satisfy the mission (technical) requirements laid out during the Conceptual Phase, but which may have different costs. As an example of this type of analysis (EXAMPLE A), the cost-effectiveness optimization of the bomb-navigation function of an aircraft system is considered. The trade-offs are assumed to be restricted to reliability, maintainability and performance. Two alternative values exist for each of these factors, representing the only possible choices between competing types of hardware and support policies. These alternatives give rise to eight competing system configurations. The required number (N) of aircraft are determined such that either of two system effectiveness criteria are met. The choice between these "equally effective" systems is then made from a comparison of the relative costs of the competing systems.

This example illustrates cost-effectiveness optimization by the exhaustion of alternatives when effectiveness is treated as a constraint (minimum acceptable value imposed from Conceptual Phase). An important point brought out by the example is the effect of a slight relaxation (critical review) of the quantitative value of the effectiveness constraint.

The principal role of models during the Acquisition Phase lies in the areas of internal optimization of the system. An early consideration in such an optimization is the allocation of the quantitative mission requirements among the subsystems of the system. Model application in this area is illustrated by reliability allocation (EXAMPLE B) and optimum payload weight allocation (EXAMPLE C).

The technique for subsystem reliability allocation considered here relates system configuration and state-of-the-art factors for each subsystem to the mission reliability constraints.

Payload allocation among three subsystems, subject to a fixed weight constraint, is treated as a two stage allocation process (dynamic programming). The first stage determines the optimum trade-off between warhead (lethal radius) and guidance (accuracy). The second stage determines the optimum division between penetration aids and the previously determined optimum mix of warhead and guidance weights. The measure of effectiveness used in this allocation is "the probability that a missile destroys a particular defended point target; given that it is delivered to the target area in a non-failed condition." Thus, this example illustrates a method for optimizing system capability, as opposed to availability or dependability.

Later in the Acquisition Phase an increasing amount of attention is devoted to questions of system support and deployment. Two examples of model use are given in the context of support optimization. The first (EXAMPLE D) treats the problem of determining the best choice of test content of a time limited prelaunch checkout. Cost considerations are introduced as a final aid in selecting between checkout formats.

The second example of support optimization (EXAMPLE E) treats the availability aspect of system effectiveness in some detail. The principal parameters of availability are defined and given explicit relation for a calendar scheduled checkout model that considers the checkout duration as system down time. Approximation formulae for the relation between maximum (optimum) availability and system parameters are derived. The questions of maintenance resource limitation and parameter estimation are briefly treated.

The last example treats the question of missile deployment (EXAMPLE F) from the point of view of a trade-off between site hardening and dispersal.

The measure of system effectiveness used here is the probability of site survival for one or more strikes against a weapon complex of several sites separated by distances of less than two lethal radii. The technique presented is partially graphical and partially analytical. The graphical portion consists of a simple square counting technique based on a transparent probability grid overlay of the geographic dispersion of the weapon complex. The probabilities determined in this way are treated as the transition probabilities which relate the weapon complex condition on a given strike to the condition which will (probably) exist on a succeeding strike. The resulting Markov chain of events is handled by matrix notation.

SECTION II

OPTIMIZATION PRINCIPLES

1.0 INTRODUCTION

The optimization of system effectiveness[1/] is a desirable goal throughout a system's life cycle, from the research and development stages through operational use. An ever-present obstacle to this goal is the problem of translating the general phrase "optimization of system effectiveness" into more precise terms. Since system effectiveness is a measure of the system's ability to accomplish its mission, it must first be determined how the system will be constituted, what its mission or objective will be, and what particular form the system effectiveness measure will assume. Associated with system effectiveness are techniques, instrumentalities, manpower, dollars, time, and other factors that determine the value of this measure. The term "resources" is used to represent these factors.

The optimization process, then, is essentially one of achieving a combination of resource use and attained effectiveness that is, by some criterion, best. This criterion is usually expressed in terms of cost, which generally represents expenditure not only of dollars, but also of time, manpower, and material. The process involves what is now commonly called cost-effectiveness analysis.

Thus, given an appropriate form for the effectiveness measure, the optimum system might be (1) one that meets or exceeds a particular value of effectiveness for minimum cost, or (2) one that attains a maximum value of effectiveness for a given total cost, or (3) some type of combination of (1) and (2) in which the costs, and possibly the effectiveness measure, can be considered as vector quantities.

After the criterion for optimization has been established, the second problem that must be faced is the representation of effectiveness as some

---

[1/] References 1 and 2 provide excellent introductions to the types of system analysis considered in this section.

function of the resources, and the resources as some function of costs. Thus, if we want to maximize effectiveness for a given cost, and if some of the resources are constrained, the mathematical form of the optimization problem would be as shown below. Let

E = system effectiveness.

C = total cost constraint

r_i = amount of i^{th} resource

$c_i(r_i)$ = cost of r_i units of i^{th} resource.

Maximize

$$E = F(r_1, r_2, \ldots r_n),$$

subject to the following constraints:

$$r_i \geq 0, \; i = 1, 2, \ldots, n$$

$$g\left[c_1(r_1), c_2(r_1), \ldots c_n(r_n)\right] \leq C$$

$$r_{i_1} \leq R_{i_1}$$

$$r_{i_2} \leq R_{i_2}$$

$$\vdots$$

$$r_{i_k} \leq R_{i_k}.$$

Note that the term resources is used in a very general sense. A resource might be the number of maintenance men available, or the electrical stresses on parts as influenced by the design or the training of operators. The function F, then, is the mathematical model that expresses effectiveness as a function of these resource variables, and the functions g and c_i represent the cost model relating these variables to total system cost.

The mathematical formulation of the optimization leads to the third problem: selecting the technique for arriving at the optimum system. The techniques range from the trial-and-error routine of examining every possible approach, to more sophisticated trial-and-error routines (such as those used in linear and dynamic programming), to extremely complicated mathematical techniques which, for complex problems, may only guarantee near or local optimum solutions.

Each of the three areas mentioned -- criterion formulation, modeling, and application of the optimization technique -- is discussed in more detail in the remainder of this section. The general approach to the optimization process is summarized in Figure 1, in which it is seen that optimization can be considered a feedback loop consisting of the following steps:

(1) "Designing" many systems that satisfy the operational requirements and constraints

(2) Computing resultant values of effectiveness and resource use.

(3) Evaluating these results and making generalizations concerning appropriate combinations of design and support factors, which are then re-fed into the model to form the feedback loop.

2.0 OPTIMIZATION CRITERION

In defining an optimizing criterion, the system analyst is faced with a problem similar to that of putting in precise, quantifiable terms the rules or criteria for choosing the "best" painting or "best" automobile. These examples do have quantifiable characteristics, such as the size of the painting or cost of the automobile; however, artistic judgment and user experience, respectively, are factors in the final choice. In the same sense, the choice of the best system is greatly influenced by the use of good engineering, economic, and operational judgment.

It is most important, however, that the optimizing criterion be defined to the maximum extent possible, for the following reasons:

(1) The inputs provided to the analyst through use of the criterion can reduce the size of the problem to a point where a judicious choice can be made.

(2) Defining a criterion forces the analyst to examine all possible alternatives in an objective manner so that the criterion can be adapted to mathematical representation and analysis.

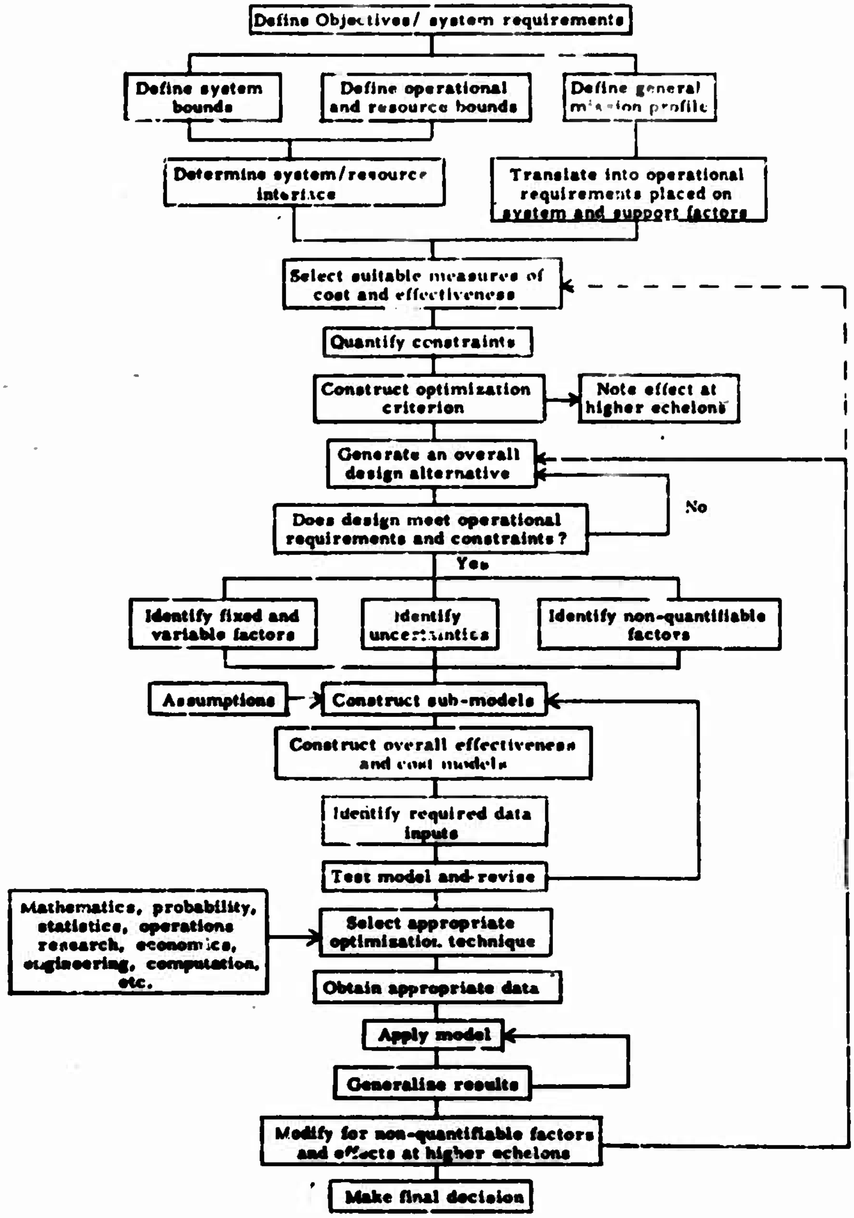


FIGURE 1. FLOW DIAGRAM FOR A GENERAL OPTIMIZATION PROCESS

(3) It is easier to incorporate the ideas and experience of others if a formal basis for optimization is established.

(4) The (partial) basis for final choice is in precise, quantifiable terms and can therefore be reviewed and revised, and can provide inputs to a learning process for future optimization problems.

When a criterion for optimization is being formulated, the system and the boundaries must be explicitly defined. This definition will influence the choice of parameters in the optimization model. The purchaser of a new automobile, for example, may or may not consider the service policies of the manufacturer and dealer. If he does, the system is both the automobile and service policies; if he does not, the system is only the automobile. In attempting to optimize a weapon system such as a bomber, the analyst has to consider whether the system is to be defined as a single bomber, a squadron of bombers, or the complete bomber fleet. It is possible that optimizing with respect to a single plane (a sub-optimization) may not yield the optimum "squadron"system, which may not, in turn give a force-wide optimum.

As part of the system-definition process, the analyst also determines the fixed and variable factors pertinent to the system. This task requires a preliminary analysis, since consideration of all possible alternatives will usually lead to problems of unmanageable size. Some factors may be considered fixed if results of previous analyses, perhaps sub-optimizations, indicate the values that have attained the best results in the past. The maintenance trouble-shooting routine, for example, might normally be considered as a variable factor, but past research in this area may be used to select a particular routine applicable to the system under study, or perhaps to restrict the range to several alternatives.

Once the mission profile is defined, consideration can be given to the physical and economic limitations that will have to be imposed. These limitations are based on requirements and availabilities, and may involve such factors as minimum system output, maximum reliability, maximum development time, maximum weight and volume, and type and number of support and operational personnel. Through such consideration and envelope of design, development, operational, and support alternatives can be established in such a way that each overall configuration within the envelope will meet physical and economic limitations as well as minimum performance goals.

Now the analyst must select a decision criterion by specifying the types of effectiveness and cost parameters to be investigated and by assigning numerical values where required. Such considerations have been treated previously, in which various types of effectiveness measures are discussed in conjunction with system and mission types. As indicated previously, the choice of objectives and criteria is perhaps the most difficult task in system effectiveness optimization. It is expected, however, that current research in the optimizing of system effectiveness will develop theory and accumulate experience to help overcome some of the difficulties of this task.

It would be impossible to establish rigid ground rules or procedures for formulating a criterion for optimizing system effectiveness. The answers to the following two basic questibns, however, will provide a great deal of insight for such formulation:

(1) Why is the system being developed?
(2) What physical and economic limitations exist?

The answer to the first question essentially defines the mission profile of the system. Where possible, the definition should be translated into quantitative parameters -- a difficult task in most cases. A performance measure such as kill-probability for a SAC bomber may be assignable, but the bomber may also have a mission to act as a deterrent -- a measure that is difficult to quantify in a completely satisfactory manner. It is for this type of multi-mission case that judgment will become especially important. Even if quantitative requirements can be placed on all mission types, weighting factors would have to be introduced to quantify the relative importance of each mission.

Factors that have relatively little impact on overall effectiveness or cost can be considered to be fixed or, possibly, can be ignored. There is, of course, a risk involved if factors chosen to be fixed or unimportant would have had a significant effect if they had been allowed to vary. Factors that fall in this "gray area" may have constraints imposed upon them in such a manner that the more detailed analysis to be performed in the optimization process will indicate final disposition. For example, if a questionable factor might have a monotonic influence on effectiveness, consideration of only extreme values might be all that is necessary to determine the significance of this influence.

It is important that factors selection, variability, and the final choice of system definition be clearly indicated so that the scope of the optimization process will be known and areas for possible modification of the formal mathematical solution will be made explicit.

3.0 MODELING

A system model is essentially a mathematical, logical, or physical representation of the interdependencies between the objectives and the resources associated with the system and its use. For dealing with the effectiveness of complex systems, the model is usually in the form of mathematical equations (mathematical model) or computer programs for simulating system operation (simulation model), or both.

On the assumption that a set of system requirements has been translated into an optimization criterion, the model builder is required, minimally, to construct a model that will enable quantification of the critical effectiveness and cost parameters as a function of the resource variables.

The overall cost-effectiveness model is usually one that consists of several sub-models, each of which may be based on models at still lower levels. Figure 2 indicates one means for sub-model classification. It should be noted that there are many other schemes for classifying models. For example, one classification scheme considers a value model, cost model, technical model, operational model, and strategic model. (See Reference 3.)

There is, naturally, a great deal of interaction among the sub-models, and model integration is required in the same sense that system integration is required.

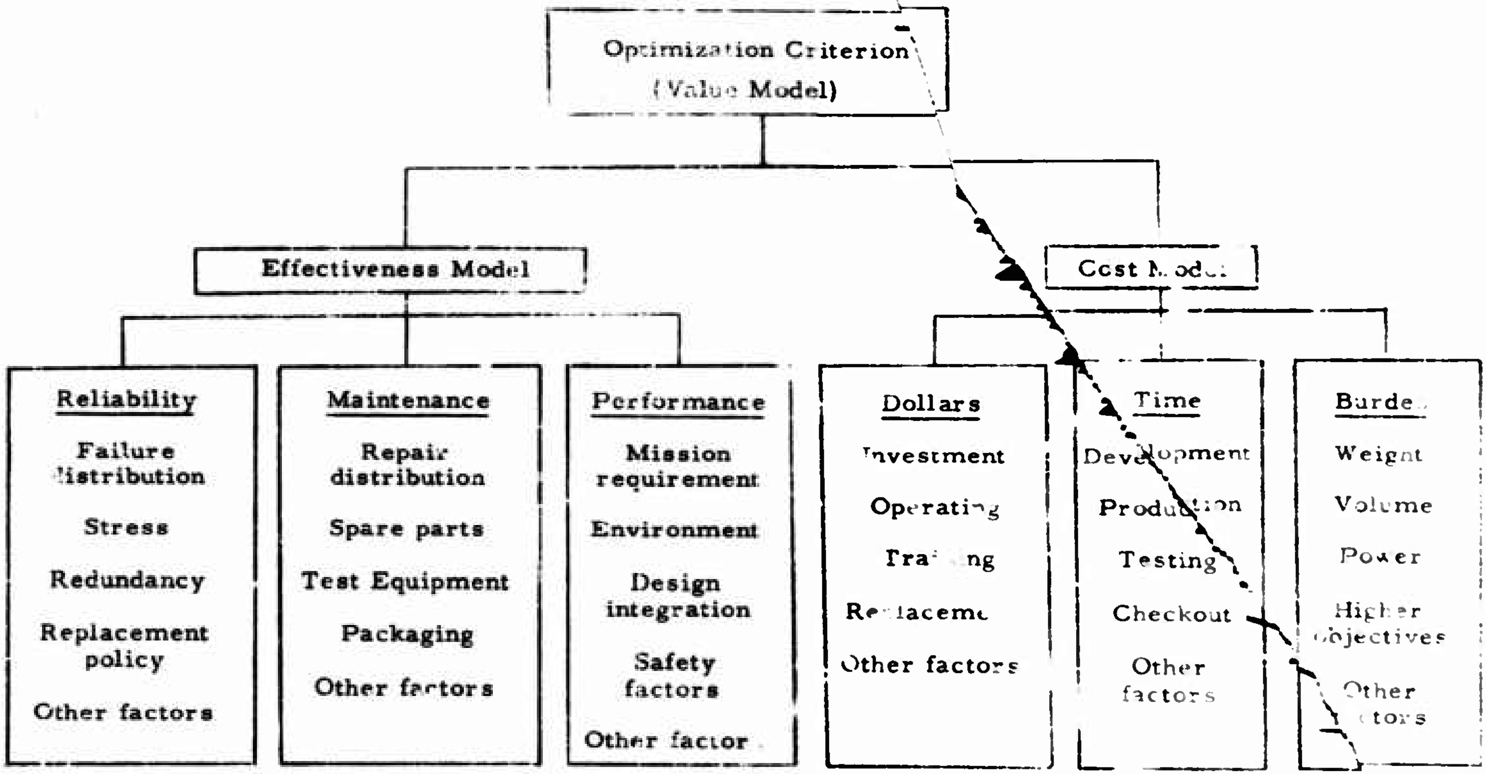


FIGURE 2

A SUB-MODEL CLASSIFICATION SCHEME

Constructing sub-models (and integrating them into an overall model) is, for most "real world" situations, still more of an art than a science, largely because the validity of the model cannot be tested through controlled experimentation. Thus the collaboration of people with wide experience in the areas of concern is an important requirement.

Sections 3.1 to 3.6 discuss some of the more important considerations for designing optimization models.

3.1 Assumptions

All assumptions required for the model should be explicitly stated and, if possible, supported by factual evidence. If no such evidence exists, it is advisable to state the reason for the assumption; e.g., mathematical simplicity or consensus of opinion, in order to indicate the degree to which the assumptions will require further justification and to pinpoint the areas in which errors might be introduced.

3.2 Adequacy

A model must be adequate in the sense that all major variables to which the solution is sensitive are quantitatively considered where possible. Many of these variables will have been preselected. Through manipulation of the model, some of the variables may be excluded or restricted, and others may be introduced. Non-quantifiable variables must be accounted for by modification of the solution rather than by direct incorporation into the model. In this sense they are quantifiable.

3.3 Representativeness

Although no model can completely duplicate the "real world," it is required that the model reasonably represent the true situation. For complex problems, this may be possible only for sub-parts of the problem, which must be pieced together through appropriate modeling techniques. As an example, analytic representation may be possible for various phases of a complex maintenance activity. The outputs from these analyses may then be used as inputs to a simulation procedure for modeling the complete maintenance process.

3.4 Uncertainty

The various types of uncertainties involved in the problem cannot be ignored, nor can they be "assumed" out; they must be faced squarely. There may be technological uncertainties involved with some of the system alternatives, operational uncertainties involved with planning and carrying out the mission, uncertainties about enemy strategy and action, and statistical uncertainties governed by the laws of chance. The simplest approach is to make "best guesses," but this may lead to disastrous results, since the probability of guessing correctly for every uncertainty is quite small. For cases involving statistical uncertainty (risk), functions-of-random-variables theory or such procedures as Monte Carlo techniques may be used. For the other types of uncertainties, the general approach is to examine all major contingencies and compute resultant cost-effectiveness parameters. The optimization criterion, then, must be adaptable for use in the evaluation of the set of cost-effectiveness results. The developments of decision theory and game theory become most applicable in the selection of a decision model in these cases, since different alternatives may be best for different contingencies.

3.5 Data

The availability of relevant data plays an important role in the development and application of a model. Data are required to support assumptions, select alternatives, and define constraints, as well as to define the cost and effectiveness constants in the proposed model. Since missing data may prevent valid model application, the model builder should investigate this possibility early in the model development stages and plan to obtain missing data or adjust the model accordingly. If a great expenditure of time and money is required to obtain the necessary data, the analyst may be forced to weigh the risks of using what is available (and making necessary assumptions) against the value received in return for the costs of the data-collection and analysis effort.

3.6 Validity

The final test of the model is whether or not it yields the best system.

Unfortunately, this determination can never be accomplished in the "real world." However, certain questions will disclose weaknesses in the model that can be corrected:

(1)	Consistency	- are results consistent when major parameters are varied?
(2)	Sensitivity	- do input-variable changes result in output changes that are consistent with expectations?
(3)	Plausibility	- are results plausible for special cases where prior information exists?
(4)	Criticality	- do minor changes in assumptions result in major changes in the results?
(5)	Workability	- does the model require inputs or computational capabilities that are not available within the research bounds?
(6)	Suitability	- is the model consistent with the objectives; i.e., will it answer the right questions?

4.0 OPTIMIZATION TECHNIQUES[2/]

As indicated in Figure 1, the technique for optimization essentially involves the application of effectiveness and cost models to all feasible designs and selection of the design which, according to the criterion, is optimum.

While this approach is conceptually simple, its implementation is virtually impossible, except for the most simple problems. Consider a problem involving fifteen variables, each of which may take one or two possible values. More than 32,000 possible system designs would have to be considered, a task that would tax even the largest of the available computers.

[2/] The differentiation between "model" and "technique" is not always clear. This is, however, primarily a semantic problem. As discussed here, a model may embody the technique for optimization, although the converse is not true.

Techniques are therefore needed to reduce the amount of mathematics and computation to a size reasonable for computer, geometrical, or even hand solution. In a sense, these techniques are sophisticated trial-and-error routines. Some of the more commonly used techniques, or fields from which such techniques are derived, are listed in Table I. The list is by no means complete. References 4 and 5 are recent publications on optimization to which the reader is referred for a description of these and other techniques and their application to selected problems.

5.0 INTERPRETATION

As indicated previously, a model of a complex process is usually incomplete because of uncertainties, non-quantitative factors, inadequate data, and inadequate consideration of the effects of the process on systems and operations at higher echelons. In such cases the results of the optimization process can only indicate the best system within the simplifications, assumptions, restrictions, and omissions required to circumvent the voids.

The effects of these circumventions must then be evaluated through some type of model feedback procedure which, on the basis of the attained results, may reveal some critical deficiencies that can be rectified.

However, even the most modern mathematical techniques and computers will yield only partial analytic solutions, mainly because of the uncertainties. These uncertainties often exist in the overall objective and, when broader contexts are being considered, it may be necessary to examine alternative objectives. We thus have the enlarged problem of first selecting the optimum mission and the associated optimum set of constraints.

The optimization process, therefore, provides the framework for a final decision. If the process is based on a correct formulation of the problem and application of a reasonable model, the decision can be critically evaluated and suitably modified. However, because of inherent limitations to a strictly analytical approach, the experience and judgment of management ordinarily inherit responsibility for the final choice.

TABLE I PARTIAL LIST OF TECHNIQUES FOR OPTIMIZATION	
I. Mathematical Techniques	Birth and death processes Calculus of finite differences Calculus of variations Gradient theory Numerical approximation methods Symbolic logic Theory of linear integrals Theory of maximum and minimum
II. Statistical Techniques	Bayesian analysis Decision theory Experimental design Information theory Method of steepest ascent Stochastic processes
III. Programming Techniques	Dynamic programming Linear programming Nonlinear programming
IV. Other Operations Research Techniques	Gaming theory Monte Carlo techniques Queuing theory Renewal theory Search theory Sensitivity testing Signal flow graphs Simulation Value theory

SECTION III

EXAMPLES RELATING TO COST-EFFECTIVENESS OPTIMIZATION

In this Section six examples are presented in detail relative to cost-effectiveness optimization, as follows:

EXAMPLE A: Aircraft System Optimization

EXAMPLE B: Reliability Allocation

EXAMPLE C: Ballistic Missile Payload Allocation

EXAMPLE D: Optimizing a Prelaunch Checkout

EXAMPLE E: Missile Availability

EXAMPLE F: A Vulnerability Model for Weapon Sites with Interdependent Elements

EXAMPLE A

AIRCRAFT SYSTEM OPTIMIZATION

ABSTRACT

A system cost-effectiveness model is developed for an Air Force training base at which daily bomber training flights are made. In the event of enemy attack, the base bomber force is assigned to targets. The objective of the example is to illustrate the optimization of the bomber effectiveness by trading off reliability, maintainability, performance and cost factors. The system effectiveness model is developed along the mathematical lines presented by Task Group II in Volume II of the final report. Optimization is accomplished by computing and comparing the costs of eight possible procurement and support policies in terms of two alternative figures of merit:

(1) For each target, there will be a 0.95 probability that at least one of the attacking aircraft will successfully accomplish the bombing run.

(2) There will be an average success probability of 0.95 for all assigned targets.

A significant aspect of this example is its illustration of the need for re-evaluating the criterion for optimization in terms of the realized output of the evaluation effort.

TABLE OF CONTENTS

ILLUSTRATIONS

TABLES

1.0 INTRODUCTION [1]

This example discusses an optimization problem that illustrates some of the concepts discussed in the main body of the report. An Air Force training base, at which daily bomber training flights are made, will be considered. In the event of enemy attack, the bomber force assigned to this base has the responsibility of attacking assigned targets. The objective is to optimize the bomber effectiveness by trading off reliability, maintainability, performance, and cost factors.

For simplicity of illustration, only the Bomb/Nav system of the bomber will be considered in the determination of bomber effectiveness. Several simplifying conditions will also be assumed for the purpose of avoiding the complex mathematical and operational procedures that may tend to obscure the objective of the example.

2.0 BASE CONDITIONS

The following operating conditions at the base are assumed:

(1) Aircraft of the allocated force are scheduled for eight-hour training flights every other day. Thus, except for grounded aircraft, half of the aircraft on the base are scheduled to fly on any given day.

(2) The takeoff times of the aircraft scheduled to fly are equally distributed over the period 0 to 1600 hours, and the landing times are equally distributed over the period 0800 to 2400 hours. It is assumed that a particular aircraft takes off and lands at the same times on every scheduled flying day.

---

[1] This example was developed jointly by ARINC Research Corporation and John Danish, Directorate of Strategic and Tactical Systems Engineering, Wright-Patterson AFB, Ohio, and originally appeared in the report "System Effectiveness: Concepts and Analytical Techniques," ARINC Research Publication 419, January 1964.

(3) At the time of scheduled takeoff, the Bomb/Nav system and the aircraft are in one of four states:

S_1	System has no malfunction; aircraft not grounded
S_2	System has undetected malfunction; aircraft not grounded
S_3	System has detected malfunction; aircraft not grounded
S_4	Aircraft is grounded.

(4) Aircraft in States S_1 and S_2 will take off on training flights as scheduled. Aircraft in States S_3 and S_4 will not take off on training flights. If an alert occurs, however, aircraft in State S_3 will be used.

3.0 OPTIMIZATION CRITERION

It is assumed that the planes under consideration are used primarily for training purposes, but that their bombing assignment (perhaps to secondary targets) in the event of enemy attack is a significant factor in optimization. Therefore, optimization will be performed with respect to actual bombing runs. Although this approach may not lead to the optimum training effectiveness, it is assumed that, for the alternatives considered, the cost and effectiveness of training vary over acceptable ranges and need not be explicitly considered unless several of the alternatives are candidates for the optimum policy.

Two optimization criteria will be considered in this example, to indicate how the choice of criteria can affect the final decision:

Criterion I - For each target, there will be a 0.95 probability that at least one of the attacking aircraft will successfully accomplish the bombing run.

Criterion II - There will be an average success probability of 0.95 for all assigned targets.

For either Criterion I or Criterion II the probability of success, which we shall call strategic effectiveness, is a function of the number of aircraft assigned to the targets and the system effectiveness of each aircraft (for this example, the Bomb/Nav system specifically). It is therefore necessary to choose the force composition, in terms of number of planes and Bomb/Nav system effectiveness, that will meet one of the criteria at minimum cost.

4.0 SYSTEM ALTERNATIVES

With only the Bomb/Nav system being considered explicitly, it is assumed that alternatives exist with respect to mission reliability and operational readiness (availability). It is further stipulated that, in case of failure of the Bomb/Nav system, a secondary mode of operation is available, but that it has a greatly reduced design capability. Since the investment cost of the planes is always greater than the cost of the reliability and maintenance alternatives, the approach will be to determine the minimum number of planes needed to meet the strategic effectiveness requirement for each alternative, and then to choose the policy with the minimum overall cost.

To simplify the problem somewhat, only a discrete number of alternatives will be considered. These alternatives are obtained from the following factors:

(1) Reliability: Two values of system failure rate are possible-- 0.10 and 0.05 failures per hour. On the assumption that these rates are constant over the mission length, t_m, the reliability alternatives are

$$R_1 = e^{-(0.10)t_m}$$

$$R_2 = e^{-(0.05)t_m}$$

(2) Maintenance: Two overall maintenance policies and procedures are possible. They involve such factors as training, administration, logistics, manpower, and equipment. They lead to two possible maintenance-time distributions, which are

identified as T_1 and T_2 in Figure 1. Maintenance time is measured from the time the aircraft lands to the time the repair action is completed.

(3) Maintenance Efficiency: The probability that a malfunction in the Bomb/Nav system is actually repaired, given a completed maintenance action, is defined as maintenance efficiency. It is a function of personnel selection, training, and test equipment, the combination of which is assumed to be largely independent of the factors affecting maintenance time. Two values are assumed:

$$M_1 = 0.6$$

$$M_2 = 0.8.$$

All possible combinations of the alternatives for the above three factors make eight possible policies, or system alternatives, available to the decision maker. They are as follows:

Policy	Alternatives	Policy	Alternatives
1	$R_1\ T_1\ M_1$	5	$R_2\ T_1\ M_1$
2	$R_1\ T_1\ M_2$	6	$R_2\ T_1\ M_2$
3	$R_1\ T_2\ M_1$	7	$R_2\ T_2\ M_1$
4	$R_1\ T_2\ M_2$	8	$R_2\ T_2\ M_2$

5.0 METHOD OF SOLUTION

By a slight modification, the basic framework for an effectiveness model described in Volume II of Task Group II can be used to determine the effectiveness of the system alternatives. In order to meet Criterion I or Criterion II, the required number of planes can be determined by a model that relates system effectiveness to strategic effectiveness. A cost model applied to each of the eight system alternatives may then be used to determine the optimum configuration.

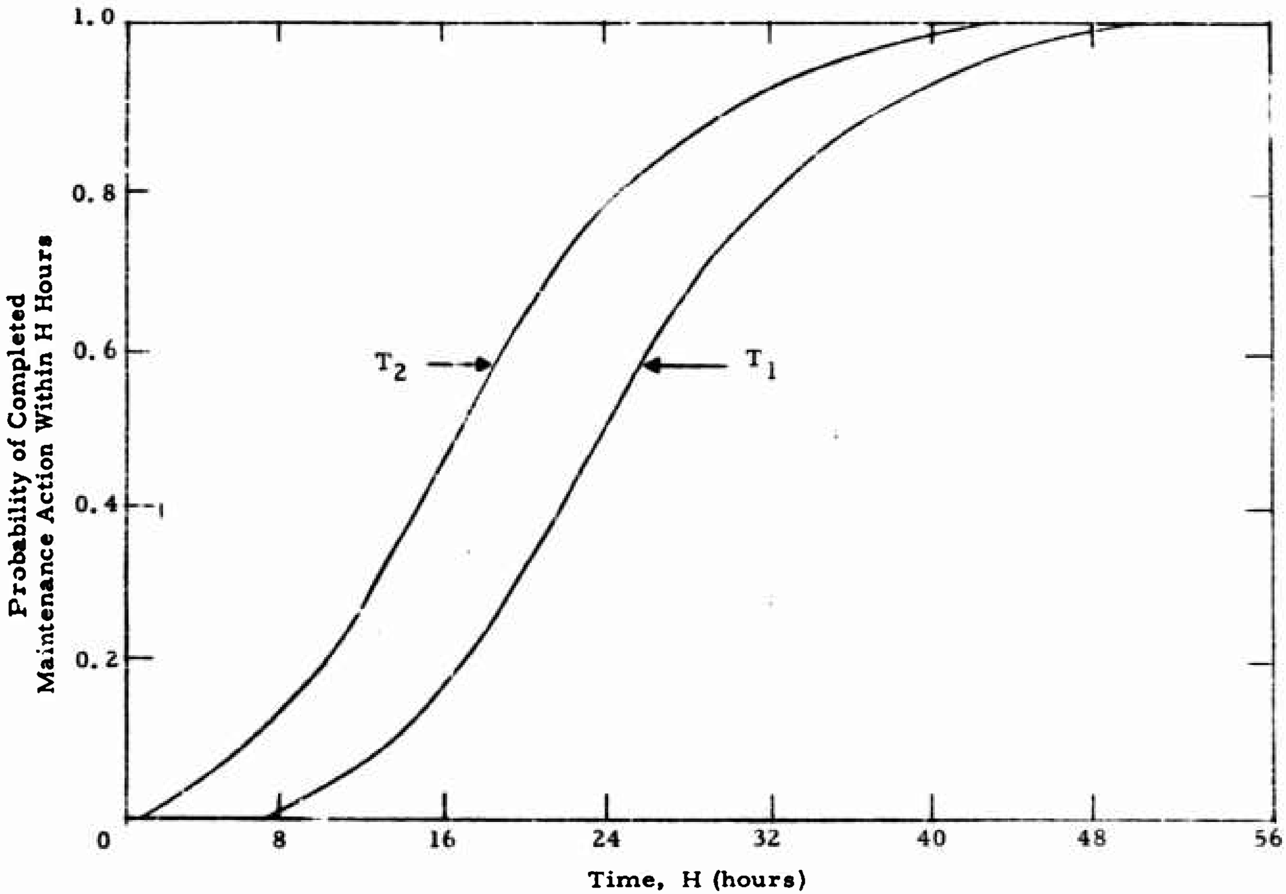


FIGURE 1

POSSIBLE DISTRIBUTIONS OF MAINTENANCE TIMES

5.1 System Effectiveness Model

5.1.1 State-Readiness (Availability) Vector

State readiness will be determined at the latest possible calendar time a plane can take off and still meet mission requirements. This time will be the time of alert, z, plus h hours, and it will be called the strike takeoff time, Z. Three system-readiness states at Z are possible:

Readiness State 1 - Aircraft is not grounded and Bomb/Nav system has no malfunctions.

Readiness State 2 - Aircraft is not grounded and Bomb/Nav system has a malfunction.

Readiness State 3 - Aircraft is grounded.

If a_i = probability, system is in State i, the state-readiness (availability) vector is

$$\overline{A}' = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$$

5.1.2 Mission-Readiness Matrix

It is assumed that a system in Readiness State 1 or 2 will be used, while a system in State 3 cannot be used. The mission-readiness matrix is then

$$\begin{bmatrix} W \end{bmatrix} = \begin{bmatrix} 1.0 & 0 & 0 \\ 0 & 1.0 & 0 \\ 0 & 0 & 0 \end{bmatrix} .$$

5.1.3 State-Transition (Dependability) Matrix

If R represents the reliability of the Bomb/Nav system for the specified mission length, the state-transition (dependability) matrix, assuming aircraft survival and no inflight repair, given Readiness State 1 or 2, is

$$\begin{bmatrix} D \end{bmatrix} = \begin{bmatrix} R & (1-R) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

5.1.4 Design-Capability Vector

If c_i represents the design adequacy of the system when it is in the i^{th} state, the design-capability vector is

$$\overline{C} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

System effectiveness is then given by

$$E = \overline{A}' \left[V\right] \left[D\right] \overline{C}$$

$$= \left[a_1 \; a_2 \; a_3\right] \begin{bmatrix} 1.0 & 0 & 0 \\ 0 & 1.0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} R & 1-R & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

$$= a_1 RC_1 + a_1 (1-R)C_2 + a_2 c_2 . \qquad (1)$$

5.2 Strategic Effectiveness Model

If a force of N aircraft is assigned to a target, the strategic effectiveness of the force, assuming probabilistic independence, is

$$S = 1 - (1-E)^N$$

for a system configuration yielding a system effectiveness value of E.

The required number of planes in a force for Criterion I is then determined from the inequality

$$N^* \geq \frac{\log (1-S_I^*)}{\log (1-E)} , \qquad (2)$$

where S_I^* is the required strategic effectiveness for each target.

For Criterion II, if S_{II}^* is the required average strategic effectiveness for k targets, the required number of planes for each target, N_i^*, is determined from the inequality

$$\frac{1}{k} \sum_{i=1}^{k} \left[1 - (1-E)^{N_i^*}\right] \geq S_{II}^* . \qquad (3)$$

5.3 Cost Model

For each policy, the cost (C) can be considered to be a function of the reliability (R), maintainability (T), maintenance efficiency (M), and required number of aircraft (N^*) associated with the policy. Hence, the cost of the i^{th} policy is

$$C \text{ (policy } i) = f_i(R, T, M, N^*), \quad i = 1, 2, \ldots\ldots, 8.$$

The details for quantifying the above models are presented in the remainder of the example.

6.0 COMPUTATIONAL DETAILS

6.1 State-Readiness (Availability)

6.1.1 Operational Sequence

To compute the state-readiness (availability) probability vector, $\overline{A}$, one must consider two cases: (1) planes scheduled to fly on training missions on the day of the alert, and (2) planes not scheduled to fly on the day of the alert. The state readiness for a bombing mission is a function of the state readiness for a training mission (S_1, S_2, S_3, or S_4).

For aircraft scheduled to fly training missions on the day of alert, Figure 2 shows the possible sequences that will lead to Readiness-State 1 (OK), 2 (F), or 3 (G) for the bombing mission at strike takeoff time, Z. It is noted that if an aircraft is flying on a training mission at the time of the alert, it is considered equivalent to a grounded aircraft. Figure 3 shows the parallel sequences for aircraft scheduled to fly a training mission on the day previous to the day of the alert.

6.1.2 Training-Mission State-Readiness (Availability)

The state probabilities at the time of scheduled training-flight takeoff are a function of the state of the system as it existed after the previous training flight, the number of maintenance hours available, the time spent on completing a maintenance action, and the maintenance efficiency. To simplify the problem it will be assumed that the probability of a plane's being grounded is a constant independent of these factors, since the factors

State at usual Takeoff Time	S_1					S_2				S_3		S_4
Schedule for Day of Alert	Training Flight									Maintenance; No Training Flight		Grounded
Occurrence of Alert	A_1	A_2	A_3			A_1	A_2	A_3				
State After Flight			OK'	F'				F'				
Maintenance Complete at Alert + h				Yes	No			Yes	No	Yes	No	
State at Alert + h	OK	Equiv. G	OK	OK	F	F	Equiv. G	OK	F	OK	F	

Legend: Occurrence of Alert
- A_1 = Before usual Takeoff Time
- A_2 = During usual Takeoff Time
- A_3 = After usual Takeoff Time
- h = Maximum Allowable Time from Alert to Strike Takeoff

State at usual Landing Times
- OK' = No Malfunction
- F' = Detected Malfunction

State at usual Takeoff Time
- S_1 = No Malfunction
- S_2 = Undetected Malfunction
- S_3 = Detected Malfunction
- S_4 = Grounded

State at Alert + h
- OK = No Malfunction
- F = Detected or Undetected Malfunction
- G = Grounded

Note: Throughout this section, whenever reference is made to a state at usual takeoff time when the alert occurs before takeoff time, the conditional phrase "if the alert had not occurred" is understood.

FIGURE 2

STATUS OF AIRCRAFT NORMALLY DUE TO FLY ON THE DAY OF THE ALERT

State at usual Takeoff Time on Day Previous to Alert	S_1			S_2		S_3		S_4
Activity on Previous Day	Training Flight					Maintenance		Grounded
State After Previous Day's Flight	OK'	F'		F'				
Maintenance Complete at Alert + h		Yes	No	Yes	No	Yes	No	
State at Alert + h	OK	OK	F	OK	F	OK	F	G

Legend: State at usual Takeoff Time on Day Previous to Alert

- S_1 = No Malfunction
- S_2 = Undetected Malfunction
- S_3 = Detected Malfunction
- S_4 = Grounded

h = Maximum Allowable Time from Alert to Strike Takeoff

State at usual Landing on Day Previous to Alert

- OK' = No Malfunction
- F' = Detected Malfunction

State of Alert + h

- OK = No Malfunction
- F = Detected or Undetected Malfunction
- G = Grounded

FIGURE 3

STATUS OF AIRCRAFT NORMALLY DUE TO FLY ON THE DAY PREVIOUS TO THE DAY OF THE ALERT pertain only to the Bomb/Nav system. A constant probability is not too unrealistic if the largest portion of grounded planes is made up of planes undergoing scheduled overhaul or preventive maintenance. An arbitrary value of 0.10 will be used for this probability.

The steady-state probabilities, P_{si}, can be obtained by the observation that transition from one state at a scheduled takeoff time to another state at the next scheduled takeoff time is a Markov process--i.e., the probability that a system is in any particular state at a scheduled takeoff time is dependent only on the state the system was in at the beginning of the previous flight, regardless of how it arrived at that previous state.

The following notation will be used:

R = Bomb/Nav reliability for an eight-hour training flight

M = maintenance efficiency, i.e., probability that a repair is actually made, given a completed maintenance action

U = probability of a completed maintenance action before the next flight, given a malfunction on landing. (U, then, is the probability of a completed maintenance action within forty hours, since takeoff and landing times are assumed constant for each aircraft.)

The transition probability matrix, given that a plane is not grounded, is then as shown in Table I. Note that P (3, 3), the probability of remaining in S_3, is zero, because if a plane was in S_3 at the beginning of the previous flight, it indicates that the maintenance was not finished, and thus the plane did not take off. Hence, up to the beginning of the present training flight, there are at least 40 + 24 + 24 = 88 hours for maintenance, and (from Figure 1) the probability of an uncompleted maintenance action (the only cause for S_3) is zero at 88 hours for both T_1 and T_2.

TABLE I

STATE TRANSITION PROBABILITY MATRIX, $[P(i, j)]$

Previous State, i	Current State, j		
	S_1	S_2	S_3
S_1	$R + \overline{R}UM$	$\overline{R}U\overline{M}$	$\overline{RU}$
S_2	UM	$U\overline{M}$	$\overline{U}$
S_3	M	$\overline{M}$	0

$\overline{X} = 1 - X$

$P(i,j)$ = transition probability from S_i to S_j

Since the Markov chain is finite and irreducible, steady-state or stationary probabilities, P_{si}, exist. These probabilities can be found from the following equations:

$$
\begin{aligned}
P'_{s1} &= P'_{s1}(R + \overline{R}UM) + P'_{s2}(UM) + P'_{s3}M \\
P'_{s2} &= P'_{s1}(\overline{R}U\overline{M}) + P'_{s2}(U\overline{M}) + P'_{s3}\overline{M} \\
P'_{s3} &= P'_{s1}(\overline{RU}) + P'_{s2}(\overline{U}).
\end{aligned}
\tag{4}
$$

These equations are subject to the condition that $P'_{si} \geq 0$, $i = 1, 2, 3$, and

$$\sum_{i=1}^{3} \quad 1.0.$$

(The primes are used to indicate the condition that the plane is not grounded.)

The solution of this system of equations is as follows:

$$
\begin{aligned}
P'_{s1} &= \frac{M}{\overline{R}(\overline{M}+\overline{U})+M} \\
P'_{s2} &= \frac{\overline{RM}}{\overline{R}(\overline{M}+\overline{U})+M} \\
P'_{s3} &= \frac{\overline{RU}}{\overline{R}(\overline{M}+\overline{U})+M}
\end{aligned}
\tag{5}
$$

If G represents the probability of a grounded plane, we then have

$$
\begin{aligned}
P_{s1} &= \bar{G}\, P'_{s1} \\
P_{s2} &= \bar{G}\, P'_{s2} \\
P_{s3} &= \bar{G}\, P'_{s3} \\
P_{s4} &= G \quad .
\end{aligned}
\tag{6}
$$

The results of applying Equations (4), (5), and (6) are shown in Table II.

TABLE II

STEADY-STATE PROBABILITIES AT SCHEDULED TAKEOFF TIME

Policy	State			
	S_1	S_2	S_3	S_4
P-1	0.617	0.226	0.057	0.100
P-2	0.746	0.103	0.051	0.100
P-3	0.646	0.237	0.018	0.100
P-4	0.777	0.107	0.016	0.100
P-5	0.706	0.155	0.039	0.100
P-6	0.801	0.066	0.033	0.100
P-7	0.728	0.160	0.012	0.100
P-8	0.822	0.068	0.010	0.100

6.1.3 Bombing Mission State Readiness (Availability)

If an alert occurs at some time, z, during the day, with a uniform probability of occurrence for $0 \leq z \leq 24$, it is necessary to find the probability that a system will be OK (a_1), F (a_2), or G (a_3) at the strike takeoff time, Z, which is h hours after the alert ($Z = z + h$). The analysis will be performed separately for the two possible groups of planes: Group A, those planes scheduled to fly training missions on the day of the alert; and Group B, those planes scheduled to fly the previous day. The sizes of these two groups are assumed to be equal.

6.1.3.1 Group A

Three mutually exclusive cases exist which relate alert time z to scheduled takeoff time t.

Case A_1: $0 \leq z < t$ (alert occurs before takeoff)
Case A_2: $t \leq z < t+8$ (alert occurs during flight time)
Case A_3: $t+8 \leq z < 24$ (alert occurs after usual landing time).

Let $f(t)$ = takeoff-time density function, and $f(z)$ = alert-time density function. Assuming uniform distributions, we have

$$f(t) = 1/16, \text{ and } f(z) = 1/24.$$

Then

$$P(A_1) = \frac{1}{384}\int_0^{16}\int_0^{t} dz\, dt = 1/3$$

$$P(A_2) = \frac{1}{384}\int_0^{16}\int_t^{t+8} dz\, dt = 1/3$$

$$P(A_3) = \frac{1}{384}\int_0^{16}\int_{t+8}^{24} dz\, dt = 1/3.$$

The following conditions are assumed:

(1) The probability that a plane is in Group A is 0.5.

(2) If case A_1 holds ($z < t$), the state at time Z is dependent on the state at the landing time (L) of the previous training flight and on the maintenance capability, if maintenance is required. As a very good approximation, we can use the following equations for this case:

$$a_1 = P_a(OK) = P_{s1}$$

$$a_2 = P_a(F) = P_{s2} + P_{s3}$$

$$a_3 = P_a(G) = P_{s4}.$$

It is noted that planes in S_2 or S_3 at t are also in these states at z for $z < t$. Planes in S_1 at t may have been in a failed state at z if the system had a malfunction at the previous landing time, and the repair was made after Z but before t. The probability that this will occur, however, is very slight since an average of 42.7 maintenance hours will have accrued by Z, and the approximation, which involves much less than 1/6 of the total planes, is therefore very good.

(3) If case A_2 holds, it is assumed that a plane cannot return to the base and be refueled and armed within h hours (h will be assumed to be equal to 1). Planes for which case A_2 holds, therefore, are equivalent to grounded planes.

(4) If case A_3 holds, the state probabilities depend on the flight reliability, on the probability distribution of the amount of time available for maintenance after landing, and on the maintenance capability.

The state probabilities at strike takeoff time (Z) for Group A planes can be determined from the equations given below. Let

a_1 = probability of completed maintenance by Z if the aircraft flies on the day of alert and the alert occurs after landing (case A_3), and

a_2 = probability of completed maintenance by Z if plane is in S_3 at t and $z > t$ (cases A_2 and A_3).

Then

$$
\begin{aligned}
P_a(OK) &= P_{s1}\left[P(A_1)+P(A_3)(R+\bar{R}V_1M)\right] + P_{s2}\left[P(A_3)V_1M\right] \\
&\quad + P_{s3}\left\{\left[1-P(A_1)\right] V_2M\right\} \\
P_a(F) &= P_{s1}\left[P(A_3)\bar{R}(1-V_1M)\right] + P_{s2}\left[P(A_1)+P(A_3)(1-V_1M)\right] \qquad (7) \\
&\quad + P_{s3}\left\{\left[1-P(A_1)\right](1-V_2M) + P(A_1)\right\} \\
P_a(G) &= P_{s1}\left[P(A_2)\right] + P_{s2}\left[P(A_2)\right] + P_{s4}\,.
\end{aligned}
$$

To compute a_1, it is necessary to obtain the conditional density of $y = Z-L$, given case A_3; and the expected probability of a completed maintenance action is, for $h = 1$,

$$a_1 = \int_1^{25} f(y\,|\,A_3)\, T(y)dy,$$

where

$f(y|A_3)$ is the conditional density of Z-L, given case A_3;

$T(y)$ is the probability that a maintenance action takes less than y hours.

Since T(y) as plotted in Figure 1 is not a standard distribution, the value of a_1 must be obtained by numerical integration, and is computed to be 0.028 for T_1 and 0.099 for T_2.

The calculation of a_2 is similar to a_1 except that, at t, forty hours of maintenance have already been accumulated. Therefore, the probability of completed maintenance by Z for cases A_2 or A_3 is high: for T_1, $a_2 = 0.969$; and for T_2, $a_2 = 0.999$.

6.1.3.2 Group B

Group B consists of planes that were scheduled for training flights on the day before the alert day. To calculate the state probabilities at strike takeoff time, let

W_1 = probability of a completed maintenance action by Z if the aircraft flew on the day before the alert (in States S_1 or S_2 at usual takeoff time on day before the alert);

W_2 = probability of a completed maintenance action by time Z if the aircraft did not fly on the day before the alert (in State S_3 at usual takeoff time on day before the alert).

Then

$$
\begin{aligned}
a_1 &= P_b(OK) = P_{s1}\left[R + \bar{R}W_1M\right] + P_{s2}\left[W_1M\right] + P_{s3}\left[W_2M\right] \\
a_2 &= P_b(F) = P_{s1}\left[\bar{R}(1-W_1M)\right] + P_{s2}\left[1-W_1M\right] + P_{s3}\left[1-W_2M\right] \qquad (8) \\
a_3 &= P_b(G) = P_{s4} .
\end{aligned}
$$

W_1 and W_2 are found in a manner similar to that used for a_1 and a_2. For T_1, $W_1 = 0.407$ and $W_2 = 0.999$; and for T_2, $W_1 = 0.644$ and $W_2 = 1.0$.

The final state probabilities at strike takeoff time are then

$$
\begin{aligned}
a_1 &= P(OK) = 1/2\left[P_a(OK) + P_b(OK)\right] \\
a_2 &= P(F) = 1/2\left[P_a(F) + P_b(F)\right] \\
a_3 &= P(G) = 1/2\left[P_a(G) + P_b(G)\right] .
\end{aligned}
\tag{9}
$$

Table III summarizes the results.

TABLE III

STATE PROBABILITIES AT STRIKE TAKEOFF TIME

Policy	State		
	OK	F	G
P-1	0.387	0.373	0.241
P-2	0.467	0.291	0.241
P-3	0.430	0.323	0.247
P-4	0.518	0.235	0.247
P-5	0.501	0.256	0.244
P-6	0.568	0.188	0.244
P-7	0.534	0.218	0.248
P-8	0.603	0.149	0.248

6.2 System and Strategic Effectiveness

With assumed values of $c_1 = 0.80$ and $c_2 = 0.25$, Equation (1) is used to obtain the Bomb/Nav system effectiveness for each policy. For Policies 1 through 4, under the assumption of a four-hour bombing mission,

$$R \text{ (4 hours)} = e^{-(0.10 \times 4)} = 0.67.$$

For Policies 5 through 8,

$$R \text{ (4 hours)} = e^{-(0.05 \times 4)} = 0.82.$$

Equation (2) can then be used to determine the number of aircraft N* that must be allocated for each target to provide a strategic effectiveness of at least 0.95 under Criterion I. Numerical values for Criterion II can be derived from the value of N*. The results of the above calculations are shown in Table IV.

6.3 Cost Estimates[2/]

The cost of each policy, or system alternative, is assumed to depend on four factors:

(1) Number of Aircraft Allocated - It is assumed that each plane costs \$9,000,000, excluding the Bomb/Nav system, and will last an average of ten years. The amortized monthly cost per aircraft is then \$75,000.

(2) Reliability - The cost for the Bomb/Nav system is amortized over a ten-year period on a single-aircraft basis as follows:

For R_1 ($\lambda = 0.10$), cost = \$2,400,000, or \$20,000 per month.

For R_2 ($\lambda = 0.05$), cost = \$3,000,000, or \$25,000 per month.

(3) Maintainability - The average monthly cost per aircraft (amortized and current) is assumed to be \$5,000 for alternative T_1 and \$15,000 for alternative T_2.

(4) Maintenance Efficiency - Total cost for alternative M_1 is assumed to be \$5,000 per month; and for M_2, \$10,000 per month. These costs are independent of the number of aircraft.

The total policy cost, C, in thousands of dollars per month for force N* is then

$$C = N^* \times \left[75 + \begin{matrix} 20 \\ \text{or} \\ 25 \end{matrix} + \begin{matrix} 5 \\ \text{or} \\ 15 \end{matrix} \right] + \begin{matrix} 5 \\ \text{or} \\ 10 \end{matrix} \quad . \tag{10}$$

6.4 Final Solutions

6.4.1 Summary Table

2/ These estimates are hypothetical and are based on simplifying assumptions.

The policy costs, along with the system effectiveness values and required numbers of aircraft, are shown in Table IV.

TABLE IV

FINAL THEORETICAL VALUES FOR EIGHT POLICIES

Policy	System Effectiveness, E	Number of Aircraft for S = 0.95, N*	Average Monthly Cost (Thousands of Dollars), C
1	0.332	7.42	747
2	0.362	6.67	677
3	0.347	7.03	779
4	0.379	6.29	702
5	0.415	5.59	592
6	0.445	5.09	544
7	0.429	5.35	620
8	0.460	4.87	570

6.4.2 Criterion I

For Criterion I it is observed (from the cost column of Table IV) that Policy 6, requiring the assignment of 5.09 aircraft to each target, will theoretically incur the lowest cost in providing a 0.95 probability of success for destroying each target. However, since only whole numbers of aircraft can be considered, it would be necessary in practice to allocate six aircraft and thereby increase the average monthly cost of this policy. In Table IV it can be observed that Policy 8 may, therefore, be the most economical, since the whole number for aircraft required (five) is only slightly in excess of the theoretical number.

When Equation (10) is used to compute the cost of applying Policy 6 with six aircraft and Policy 8 with five aircraft, the above suggestion is found to be correct. (See Table V.)

TABLE V PRACTICAL FORCE AND COST VALUES FOR POLICIES 6 AND 8		
Policy	$N^{*\prime}$ | $S^* \geq 0.95$	Cost C'
6	6	640
8	5	585
Primes indicate Practical Values		

Since strategic effectiveness under Criterion I is defined as the probability that each target in a proposed multi-target strike will be attacked successfully, the required size of the force is simply the appropriate multiple of N*. Therefore, Policy 8 is optimum not only for a one-target strike force but also for strike forces designated for any number of targets. Figure 4 presents the average monthly costs, under Policies 6 and 8, of the forces required for potential strikes on 1 to 20 targets, under the stipulation of Criterion I.

Several questions arise concerning Criterion I. While it is true that Criterion I strictly requires at least a 0.95 strategic effectiveness for each target, Policy 6 under the practical case yields a value of S = 0.971, and Policy 8 yields a value of S = 0.954. Thus, while Policy 8 is less costly than Policy 6, the latter yields a higher value of S, although Policy 8 is still better in terms of "strategic effectiveness per dollar."

One may also consider the effects of being "slightly below" the 0.95 requirement. Since Policy 6 required a theoretical 5.09 planes per target to yield a strategic effectiveness of 0.95, it would seem that if five planes were allocated, the value of S would be quite close to 0.95 and, in fact, does equal 0.947. This value would be achieved at an average monthly cost of $535,000. In terms of effectiveness per hundred thousand dollars, Policy 6, with five planes, yields a value of 1.77, while Policy 8, with five planes, yields a value of 1.63.

<u>These questions emphasize the need for re-evaluating the criterion</u>

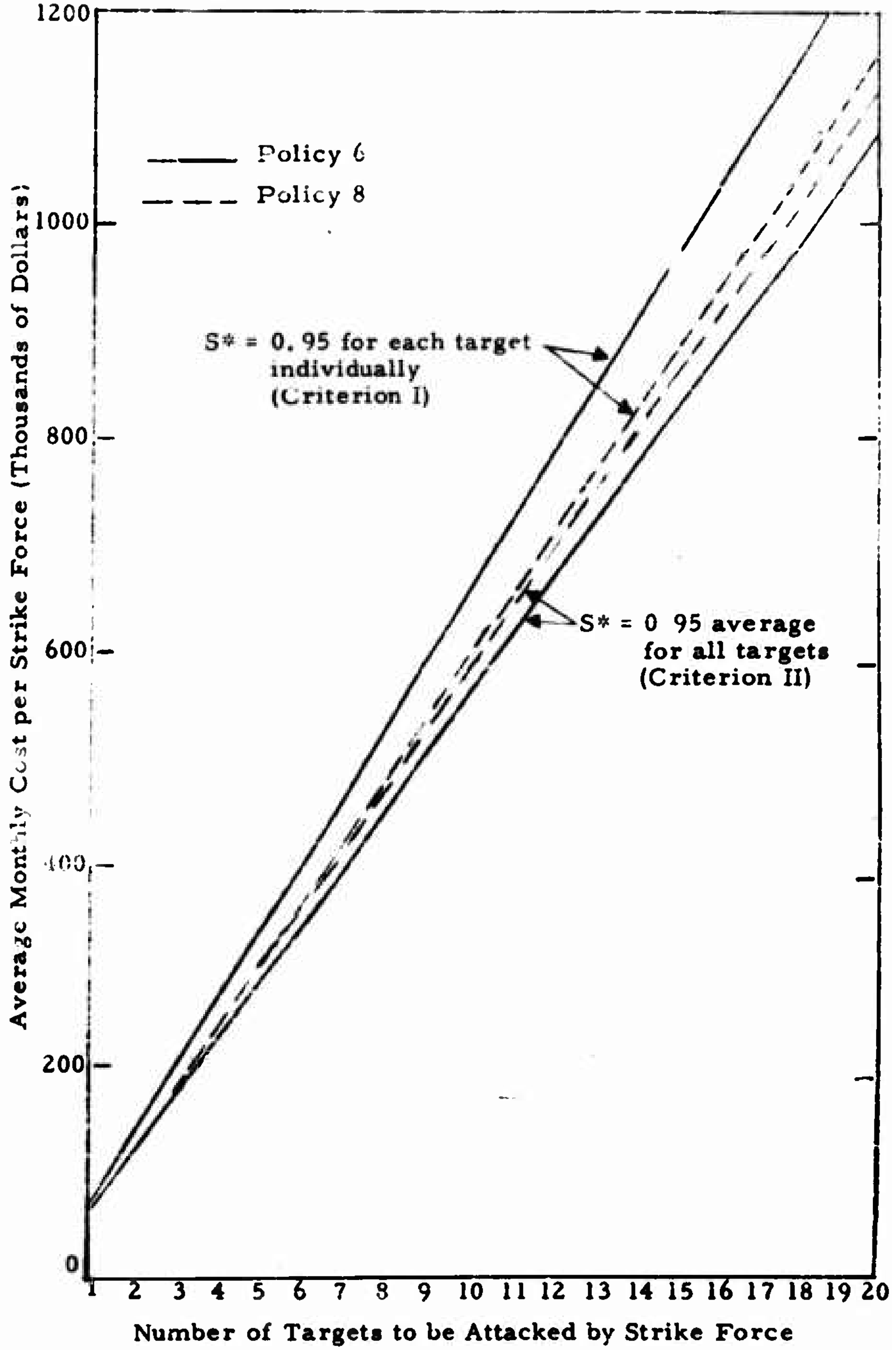


FIGURE 4

AVERAGE MONTHLY COST FOR VARIOUS STRIKE FORCES for optimization by some iterative process. For this example, the choice between Policies 6 and 8 is not clear; perhaps other factors, such as training effectiveness and policy feasibility, must be introduced.

6.4.3 Criterion II

Criterion II stipulates an average strategic effectiveness of 0.95 across all targets. For a two-target strike under this stipulation, for example, only 11 aircraft would be allocated under Policy 6, instead of the 12 previously required under Criterion I. A value of S = 0.971 would be provided for one target by a force of six aircraft, while S = 0.947 would be provided for the second target by a force of five aircraft. The average effectiveness, assuming equal target priority, is greater than 0.95 and thus satisfies the requirement.

Figure 4 also presents the average monthly costs under Policies 6 and 8 for the forces required for potential strikes on 1 to 20 targets when the requirement is $S^*_{II} \geq 0.95$ across all targets. It may be observed that for strikes involving one or two targets Policy 8 is optimum, but for three or more targets Policy 6 is optimum.

The results of considering departures from the required minimum of $S^*_{II} = 0.95$ are presented in Figure 5 which shows the values of S_{II} achieved by Policies 6 and 8 for a potential strike involving 1 to 20 targets when the force size ensures that $S'_{II} \geq 0.95$. It may be conjectured from the figure that oscillations will dampen out to some value slightly above 0.95 as the number of targets increases beyond 20.

The strategic effectiveness values shown in Figure 4 were divided by the average monthly cost, computed under both policies for each of the 20 targets, to construct Figure 6. This figure shows that there is some basis for favoring Policy 6 over Policy 8 for multiple targets in terms of "strategic effectiveness per dollar;" but, as with Criterion I, the policy choice may better be made after the inclusion of other factors.

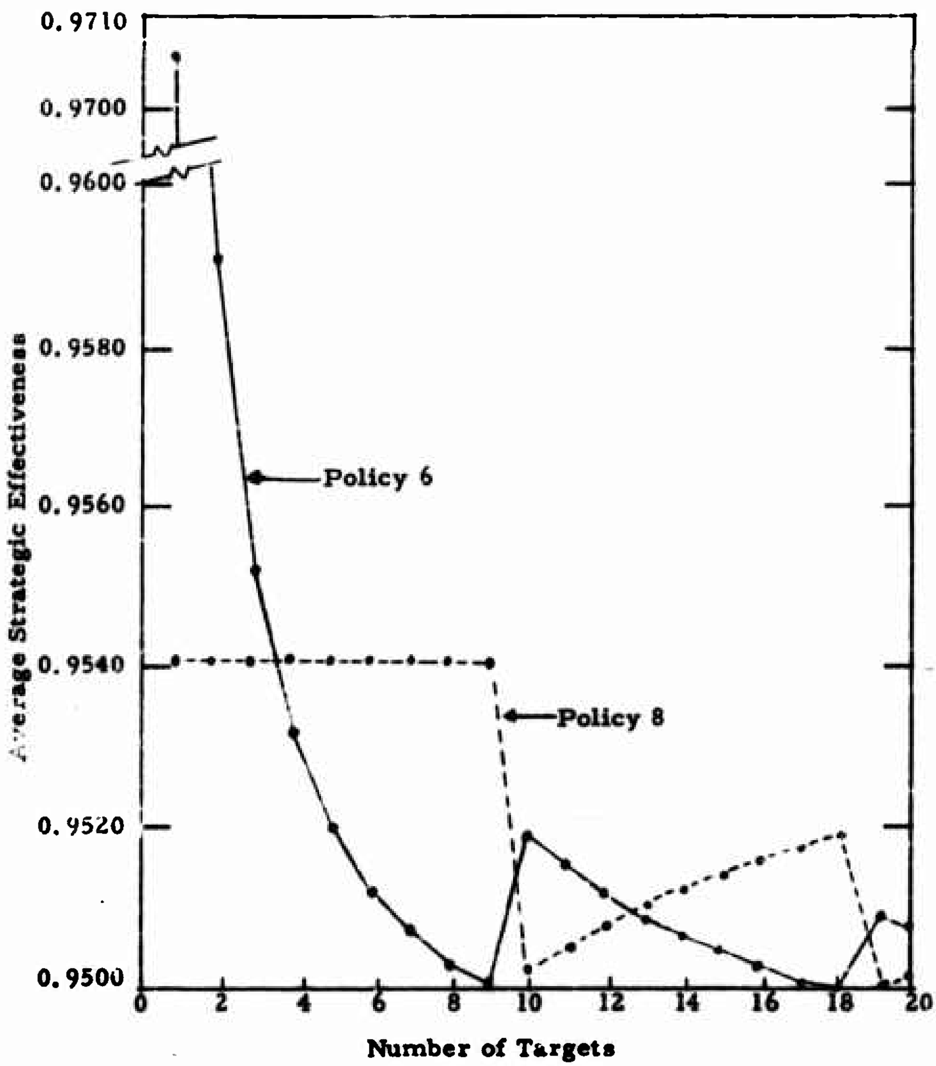


FIGURE 5

AVERAGE STRATEGIC EFFECTIVENESS ≥0.95 WITH MINIMUM COST

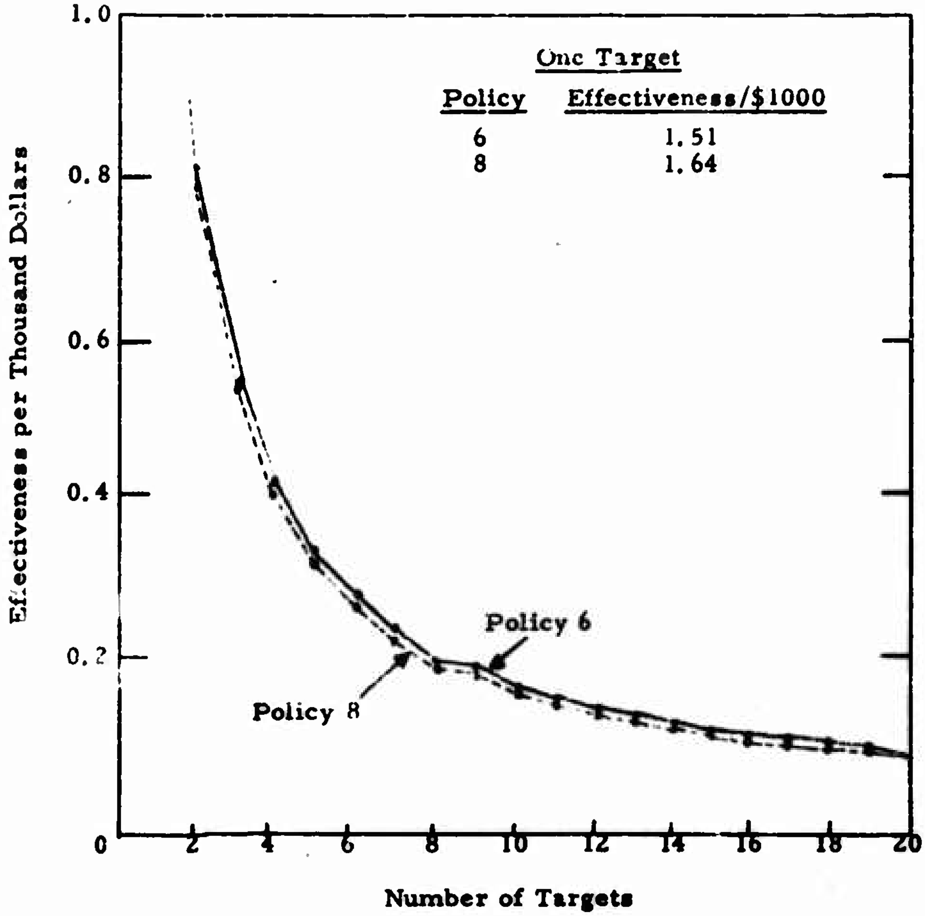


FIGURE 6

EFFECTIVENESS PER DOLLAR FOR MINIMUM COST, AVERAGE STRATEGIC EFFECTIVENESS > 0.95

EXAMPLE B

RELIABILITY ALLOCATION

ABSTRACT

A method for allocating system reliability requirements among subsystems (or lower level units) is presented. The method considers serial and redundant interconnections among the subsystems. The relationship between system reliability requirements and system effectiveness requirements is considered.

TABLE OF CONTENTS

ILLUSTRATIONS

1.0 INTRODUCTION

Military agencies responsible for providing weapon systems for operational use must translate the overall system requirement into quantitative reliability requirements at many system sub-levels. Realistic and consistent reliability requirements for units, equipments, or subsystems must be assigned in order to achieve and demonstrate specified operational weapon-system reliability. This assignment is commonly called "reliability allocation."

The allocation process is an important element in an overall optimization program since it essentially provides a set of optimum reliability goals based on a selected decision criterion. At a minimum, it provides a set of requirements on reliability characteristics which represent achievable goals in terms of the state-of-the-art. One can, however, also include in the allocation model such factors as maintenance, cost of reliability improvement, and cost, weight and space factors associated with redundancy in order to obtain a set of goals or requirements that are not only representative of current capabilities but involve variables which "act" on these capabilities.

The allocation method presented here is based primarily on state-of-the-art factors, system configuration and mission requirements. Step-by-step procedures for implementing the method are presented.

2.0 FACTORS INCLUDED IN THE ALLOCATION METHOD

This section describes those factors and influences which are considered to be of sufficient importance for explicit inclusion in the method. These factors, therefore, will have to be translated to quantitative terms or be capable of mathematical representation and analysis.

2.1 System and Failure Definitions

The system under consideration must be clearly defined in terms of its functions and boundaries. The conditions that constitute failure or unsatisfactory performance can be determined from a study of the operational demands and the functional requirements of the system. Those conditions can then be translated into measurable unit characteristics. The boundaries surrounding the system and each unit must be clearly defined to insure that important items are neither neglected nor considered more than once.

2.2 System Reliability Requirement

The primary element in a reliability allocation method is the system reliability requirement. It is usually determined on the basis of ultimate user requirements and feasibility, but it may derive from an allocation performed at a higher echelon. The requirement may be stated in any appropriate measure such as mean life, system failure rate, or preferably, a reliability requirement for a specific period of time.

The success-probability requirement on a weapon system may be based on the desires of field personnel who, naturally, think in terms of the probability that the system can successfully complete some specific mission, probably under wartime conditions. The supplier of the system cannot, however, design or test the system under these same conditions. The translation, therefore, must be made in the writing or interpretation of a specification, which requires certain measurable system and equipment parameters to be within specified limits under specified environmental conditions, with the implication that hardware meeting these requirements will also fulfill the military mission. 1/

We shall ordinarily interpret a success-probability requirement in the usual reliability sense. That is, a requirement of $R^*(T)$ shall be interpreted to be the probability that the system will satisfactorily meet all design specifications for T hours of operation under stated conditions. By

---

1/ As part of the allocation process, one should first investigate the feasibility of the overall requirement. This may be done by considering past reliability performance of systems of similar complexity operating in a similar environment. Feasibility determination is not discussed here.

assuming values for operational readiness (availability) and design capability, one can obtain the equivalent reliability requirement from a requirement stated in terms of a system effectiveness parameter.

2.3 Unit State-of-the-Art

State-of-the-art measures are required in order to determine the relative reliabilities of the allocation units within the system. These are the basic data inputs in a typical reliability allocation procedure and are usually stated in terms relatable to the measure used for the system reliability requirement. Relative average failure rates are the state-of-the-art measures adopted in this study. They will give exact answers for units with constant failure rates; furthermore, they represent a reasonable approach for other typical failure densities.

2.4 Relationship Between Unit and System Failure

The relationships between unit failure and system failure must be determined before the allocation is made. Four types of basic relationships, for which allocation methods are presented, are as follows:

(1) Serial system: no functional duplicates exist and each unit must operate successfully for system success.

(2) Modified serial system: no functional duplicates exist but units can fail without necessarily causing system failure.

(3) Redundant system: components of the system are duplicated for increased reliability but each redundant path or mode of operation is equally effective in performing its function.

(4) Multimodal system: redundant paths or modes of operation are not equally effective in performing their function. [2/]

[2/] Multimodal systems are not considered here. An approach for reliability allocation of multimodal systems is presented in the report, The Allocation of System Reliability, H. Balaban and H. Jeffers, ARINC Research Corp., Publication No. 274, Volume I, November 1961.

2.4.1 Unit Essentiality

The concept of essentiality, used to describe the effects of unit failure on mission success, is considered only if a failed unit has no functional duplicate. It is defined as follows.

> The essentiality of a unit is the probability that the system will fail to accomplish its mission if the unit fails while all other units perform satisfactorily.

An example of a unit which might have an essentiality less than one is a radar beacon transmitter on a satellite used for tracking purposes. If the beacon fails after the orbit has been firmly established, the orbital position may possibly be obtained through mathematical analysis.

Unit essentiality must be considered in the allocation of reliability of modified serial systems; it may also be involved in redundant and multi-modal systems. At the design stage of system development, the likelihood is that the essentiality of various units within the system will have to be assigned intuitively on the basis of experience gained with similar systems. If appropriate system failure data is available, essentiality can be estimated by the ratio,

$$E_j = \frac{\text{Number of Mission Failures due only to } j^{th} \text{ Unit Failure}}{\text{Number of } j^{th} \text{ Unit Failures}}$$

2.5 Unit Duty Cycles

Duty cycles must be included in an allocation method to reflect any variance in unit operational time requirements with respect to systems operation-time. Units which have a limited operational period because of a low duty cycle (e.g., the hydraulic system of an airplane) should have a relatively high allocation over the system operating period.

3.0 MATHEMATICAL MODEL FOR ALLOCATION METHOD

3.1 Requirements of the Allocation Model [3/]

The following requirements exist in developing a reliability allocation model:

(1) Allocated unit reliability increases as unit state-of-the-art decreases.

(2) Allocated unit reliability increases as essentiality increases.

(3) Allocated unit reliability increases as duty cycle or required time of operation decreases.

(4) Units in a system with equal essentiality, duty cycle and state-of-the-art should have the same allocated reliability whether in series or in a redundant configuration within the same system.

3.2 Assumptions

The following two basic assumptions are made in developing the allocation model:

(1) Allocation units can be so chosen that failure probabilities are independent.

(2) Unit state can be described in discrete terms of success and failure.

These two assumptions greatly simplify the mathematics of allocation and are believed to be reasonable for the purposes of a design-stage reliability allocation method. With regard to the first assumption, if components within the system are known to be dependent, they may possibly be grouped into one allocation unit, making the failure probability of this unit independent of the state of other units. The state-of-the-art of this unit can then be adjusted for the dependence that exists.

The second assumption is reasonable in the sense that reliability, by definition, requires that satisfactory performance be uniquely defined.

3/ Cost is not treated as a factor here.

In practice, this is often a most difficult problem, and success/failure definitions are necessarily somewhat arbitrary. If allocation is primarily regarded as a procedure for defining parameters of reliability acceptance tests, such tests usually require that this second assumption be satisfied in order to determine if a unit has passed. In this case, the same success/failure definitions for such tests should hold for allocation. The allocation model does not require explicit success/failure definitions, but since the input data is based on success/failure appraisals of field personnel, it is implicitly assumed that similar appraisals can be made for the units under consideration.

3.3 Data Input

The data inputs to the model which reflect the unit state-of-the-art are called the unit failure indices. These indices are obtained from the relative failure rates of the elements in a unit which are based on such factors as type of active element, type of function, part types, and environment.

If the j^{th} unit has a failure index of K_j, the total system failure index for a serial system [4/], K, is defined by

$$K = \sum_{j=1}^{n} K_j \quad . \tag{1}$$

The failure index ratio or relative weight of each unit is

$$w_j = \frac{K_j}{K} \quad . \tag{2}$$

The basis for the allocation is that each unit of w has an equal effect on system reliability in the same sense that each unit of failure rate has an equal effect on reliability. Thus, the failure contribution of the j^{th} unit is proportional to w_j and as shown in Section 3.4, the effect can be quantified by using the w_j as exponent weighting factors.

4/ See Section 3.6 for a discussion of redundant systems.

3.4 Allocation Method for Serial Systems

In this section we shall derive the allocation equation for a serial system consisting of n independent units. From basic reliability theory, for a system of this type, the system reliability function is

$$R(T) = \exp\left[-\sum_{j=1}^{n} \int_0^{t_j} Z(\tau_j)d\tau_j\right] \tag{3}$$

where

$R(T)$ is the system reliability over T system hours

t_j is the required operating time of the j^{th} unit over T system hours

$Z(\tau_j)$ is the hazard rate function for the j^{th} unit.

For the exponential distribution, $Z(\tau_j)$ is constant over time and is equal to what is commonly called the failure rate, λ_j. For other failure densities, if the hazard rate function is approximately constant over t_j hours, e.g., the effects of wearout in time period $(0, t_j)$ are negligible, the $Z(\tau_j)$ in equation (3) can be replaced by average hazard rates estimated from appropriate failure data. A common formula for obtaining such an estimate is

$$\lambda_j = \frac{\text{Number of failures in } (0, t_j)}{\text{Total accumulated operating time}} .$$

The substitution of λ_j for $Z(\tau_j)$ in equation (3) yields

$$\begin{aligned} R(T) &= \exp\left[-\sum_{j=1}^{n} \lambda_j t_j\right] \\ &= \exp\left[\lambda_s T\right] \end{aligned} \tag{4}$$

where

λ_s is the estimated system hazard rate for $(0 < \tau < T)$

λ_j is the estimated j^{th} unit hazard rate for $(0 < \tau_j < t_j)$.

If a system reliability requirement of $R^*(T)$ exists, one can find an equivalent λ^*_s by equation (4). Since $\lambda_s T = \sum_{j=1}^{n} \lambda_j t_j$, allocated average unit failure rates of $\hat{\lambda}_j$ must be determined so that

$$\sum_{j=1}^{n} \hat{\lambda}_j t_j = \lambda^*_s T \quad . \tag{5}$$

A reasonable approach is to equate each $\lambda_j t_j$ in equation (5) to $w_j \lambda^*_s T$ since the reliability contribution of the jth unit is proportional to w_j. Hence, allocated average unit failure rates can be determined from the equation.

$$\hat{\lambda} t_j = w_j \lambda^*_s T \tag{6}$$

or

$$R^*(T) = e^{-\lambda^*_s T}$$

$$= e^{-w_1 \lambda^*_s T} \; e^{-w_2 \lambda^*_s T} \ldots \; e^{-w_n \lambda^*_s T} \tag{7}$$

The allocated reliability of the j^{th} unit for t_j operating hours is

$$\hat{R}(t_j) = e^{-w_j \lambda^* T} = \left[R^*(T)\right]^{w_j} \tag{8}$$

Allocated failure rate requirements are obtained as follows:

$$\hat{\lambda}_j = \frac{w_j \lambda^*_s T}{t_j} = -\log \hat{R}(t_j)/t_j. \tag{9}$$

For $R(t_j) \geq 0.9$,

$$\hat{\lambda}_j \approx -\frac{w_j}{t_j} \log R^*(T). \tag{10}$$

Mean life requirements can be obtained as the reciprocal of the allocated failure rates.

An average failure rate for units in which the hazard rate function is not constant can be allocated by the equation

$$\hat{\lambda}_j = \frac{1-\hat{R}(t_j)}{t_j} \quad . \tag{11}$$

3.5 Allocation Method for Modified Serial Systems

For modified serial systems, one or more units have essentialities less than one and therefore these units may fail without necessarily causing system failure. The probability that the system will not fail due to failure of the j^{th} unit is

$$1 - E_j \left[1 - R(t_j)\right] \tag{12}$$

Under the assumption of independent unit failures and serial operation, a good approximate formula for system reliability is

$$R(T) = \prod_{j=1}^{n} \left\{1 - E_j \left[1 - R(t_j)\right]\right\} \tag{13}$$

This formula is approximate in the sense that it implies independence of unit essentialities, e.g., the probability of system failure given failure of units A and B is $E_a\ E_b$. Since E will most likely be one for the majority of units, the above equation is reasonable.

If $R^*(T)$ is the system reliability requirement, the allocated contribution of the j^{th} unit to system reliability as given by equation (8) is

$$\left[R^*(T)\right]^{w_j}$$

Hence, by equation (12), $\hat{R}(t_j)$ must be chosen so that

$$1 - E_j \left[1 - \hat{R}(t_j)\right] = \left[R^*(T)\right]^{w_j}$$

or

$$\hat{R}(t_j) = 1 - \frac{1 - R^*(T)^{w_j}}{E_j} \tag{14}$$

This is the formula derived in the AGREE report [5/]. It is important to note that E_j must be greater than $1 - R^*(T)^{w_j}$ in order to avoid negative reliability allocations. In most practical situations, especially where

---

5/ Advisory Group on Reliability of Electronic Equipment, Office of the Assistant Secretary of Defense, Reliability of Military Electronic Equipment, June 4, 1957, pp. 52-57.

weight and space are at a premium, units with low essentiality and high failure indices are not common. If a unit does exist for which $E_j > 1 - R^*(T)^{w_j}$, it is recommended that this unit be eliminated from the allocation and the failure index ratios, w, of the remaining units be recomputed.

Failure rate and mean life allocation equations can be derived in the same manner as for serial systems. By computing $\hat{R}(t_j)$ from equation (14), equations (9) and (10) remain unchanged. The approximate formulas for failure rate and mean life allocations become

$$
\begin{aligned}
\hat{\lambda}_j &= - \frac{E_j t_j}{w_j \log R^*(T)} \\
\hat{\theta}_j &= - \frac{w_j \log R^*(T)}{E_j t_j}
\end{aligned}
\tag{15}
$$

3.6 Allocation Method for Redundant Systems

A redundant system is defined here as one in which some (or possibly all) of the elements have functional duplicates for purposes of increasing system reliability. Each redundant path or mode of operation is assumed to be equally effective in performing its function, i.e., the design capabilities of all modes of operation are equal.

Two specific redundancy types are considered:

(a) Active-parallel or continuous redundancy where all redundant units are continuously energized at any one time.

(b) Standby or sequential redundancy where only one of the redundant units is energized at any one time.

If switching is involved (as it always is for standby redundancy), the probability of premature switching (switching when not required) shall be assumed to be relatively small as compared to the probability of failure to switch when required. The switching mechanism, if it is subject to failure, can therefore be considered as a series unit.

The following model applies only to redundant systems which contain a single redundant configuration, i.e., only one unit or one group of units is duplicated. The degree of redundancy is fixed at two, i.e., there are only two paths of operation for the particular function which is duplicated. The latter restriction was made primarily because of the belief that, at the design stage, redundancy is not and should not be used extensively since the technique can be employed much more effectively after allocations are made and predictions or laboratory tests performed to determine possible trouble areas. The extension of the model to degrees greater than two is easily made and briefly discussed in Section 3.6.1.

The restriction on the number of redundant configurations is also justified by the above argument and, in addition, the complexity of the allocation model is greatly increased for more than one configuration. If the system has two or more redundant configurations, an approximation that will yield conservative allocations can be made.

Assume two units are duplicated in a redundant design. The reliability block diagram of the system will therefore be as shown in Figure 1.

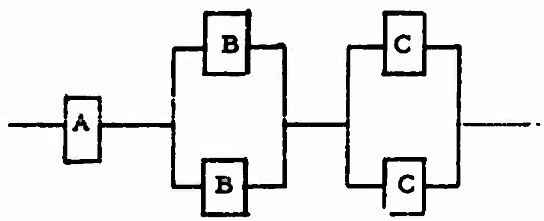


FIGURE 1

BLOCK DIAGRAM OF SYSTEM WITH REDUNDANCY

"A" represents all series units. By eliminating the cross-connects between the B and C configurations, the block diagram reduces to that shown in Figure 2.

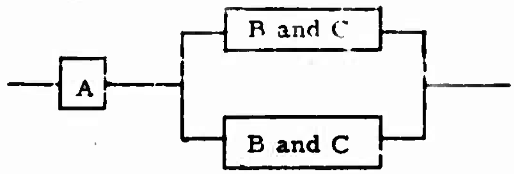


FIGURE 2

BLOCK DIAGRAM EQUIVALENT TO FIGURE 1

This is a single redundant configuration for which the method applies. The method also permits allocation to the individual B and C units as well as to the redundant configuration and to the redundant units composed of B and C. Since the reliability of the second system is generally lower than that of the first, the reliabilities allocated will be somewhat higher than actually required.

3.6.1 Identical Redundant Paths

The reliability block diagram of a system with a single redundant configuration consisting of identical redundant paths is shown in Figure 3.

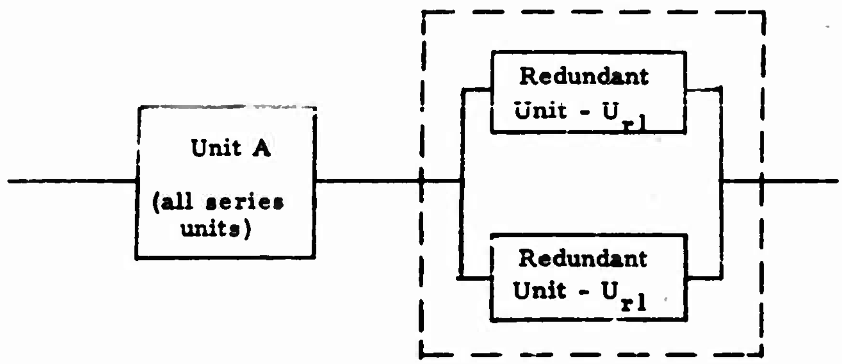


FIGURE 3

REDUNDANT CONFIGURATION - U_r

(Identical Redundant Units)

Unit A represents the combination of all series units. Unit U_{r1} is a redundant unit (possibly including more than one allocation unit) which is duplicated to form the redundant configuration. K_a shall be used to designate the total failure index of the series unit, and K_{r1} the total failure index of each redundant unit.

The method used to allocate the system reliability requirement to Unit A, to the redundant configuration, and to each redundant unit is to determine an equivalent complexity for the redundant configuration K_r which will justify use of the basic allocation formulas for serial or modified serial systems. The derivation for determining K_r is given below.

Equation (8) gives the basic allocation equation for serial systems

$$\hat{R}_j = R^{*w_j} \quad \text{or} \quad w_j = \frac{\log \hat{R}_j}{\log R^*}$$

(operating times can be neglected for the present).

Since redundant configuration U_r is in series with Unit A, allocations based on equation (8) can be performed if values can be found for w_a, the failure index ratio of U_a, and for w_r, the failure index ratio of U_r. By definition,

$$w_a = \frac{K_a}{K_a + K_r} \qquad\qquad (16)$$

$$w_r = \frac{K_r}{K_a + K_r}$$

where K_r is as yet undetermined.

If some sub-combination of units in Unit A (the series unit) had a total failure index of K_{r1} (the failure index of each redundant element), the failure index ratio of this sub-combination is

$$w_{r1} = \frac{K_{r1}}{K_a + K_r} \qquad\qquad (17)$$

Since units with the same failure index ratio are required to have the same allocated reliability (assuming equal essentiality and duty cycle) whether in series, parallel, or both, w_{r1} above is also the failure index ratio of U_{r1} and U_{r2}. By equations (8), (16), and (17)

$$\frac{w_r}{w_{r1}} = \frac{\log \hat{R}_r}{\log \hat{R}_{r1}} = \frac{K_r}{K_{r1}}$$

Hence

$$K_r = \frac{K_{r1} \log \hat{R}_r}{\log \hat{R}_{r1}}$$

Substituting for K_r in equation (16) yields

$$w_r = \frac{K_{r1} \log \hat{R}_r}{K_a \log \hat{R}_{r1} + K_{r1} \log \hat{R}_r} \tag{18}$$

Since we also have

$$w_r = \frac{\log \hat{R}_r}{\log R^*}$$

equation (18) can be rewritten and simplified to

$$\log \hat{R}_r = \frac{K_{r1} \log R^* - K_a \log \hat{R}_{r1}}{K_{r1}}$$

or

$$\log \hat{R}_r = \log R^* - \alpha \log \hat{R}_{r1} \tag{19}$$

where

$$\alpha = \frac{K_a}{K_{r1}} \tag{20}$$

In general, $\hat{R}_r$ is some function of $\hat{R}_{r1}$, the allocated reliability of the redundant units, e.g., for active-parallel redundancy

$$\hat{R}_r = 2\hat{R}_{r1} - (\hat{R}_{r1})^2 .$$

Hence, by the inverse relationship,

$$\hat{R}_{r1} = 1 - (1 - \hat{R}_r)^{1/2}$$

(Note: Since relationships of this type exist for any number of redundant units, the method applies to all degrees of redundancy.)

Writing $\hat{R}_{r1}$ as some function, $\hat{R}_{r1} = f(\hat{R}_r)$, we have from equation (19)

$$\log \hat{R}_r = \log R^* - \alpha \log [f(\hat{R}_r)] \quad . \tag{21}$$

For a given α and R^*, equation (21) can be used to determine $\hat{R}_r$ for a specific type of redundancy. The remainder of the system (Unit A) is then allocated a reliability of $R^*/\hat{R}_r$. It is possible, however, to use equation (21) to determine K_r directly as shown below.

From equation (16)

$$K_r = \left(\frac{w_r}{1 - w_r}\right) K_a \tag{22}$$

By equation (8)

$$w_r = \frac{\log \hat{R}_r}{\log R^*}$$

Hence, for a given α and R^*, equation (21) can be used to obtain $w_r(\alpha, R^*)$. This enables us to obtain the ratio

$$Z(\alpha\, , \ R^*) = \frac{w_r(\alpha\, , \ R^*)}{1 - w_r(\alpha\, , \ R^*)} \tag{23}$$

Then from equation (22)

$$K_r = Z(\alpha\, , \ R^*)\, K_a \tag{24}$$

(Nomographs giving values of $Z(\alpha, R^*)$ for wide ranges of α and R^* for both active-parallel and standby redundancy are presented in Figures 8 to 11 of Section 4.0). Once K_r is determined, the total failure index of the system can be found by

$$K = K_1 + K_2 + \ldots\ldots + K_m + K_r$$

where K_1 to K_m are the failure indices of the units in series (represented by K_a in the above derivation). Failure index ratios are then found by

$$w_j = K_j/K$$

for each series unit, the redundant configuration and for each redundant unit as well. The allocation equations for serial or modified serial systems then apply.

3.6.2 Duplicate Systems

For designs where the complete system is duplicated, allocation is relatively simple. For active-parallel operation, the reliability requirement for each system is

$$\hat{R}_{r1}(T) = 1 - [1 - R^*(T)]^{1/2}$$

$\hat{R}_{r1}(T)$ can then be considered to define $R_{r1}^*(T)$, the reliability requirement of each system, which then can be sub-allocated among the units of the system by methods previously described. For standby redundancy (under the assumption of constant failure rates),

$$R^*(T) = \hat{R}_{r1}(T) \; [1 - \log \hat{R}_{r1}(T)]$$

For a given $R^*(T)$, $\hat{R}_{r1}(T)$ can be graphically determined and sub-allocations within the system for $R_{r1}^*(T) = \hat{R}_{r1}(T)$ can be performed.

3.6.3 Dissimilar Redundant Paths

Assume the block diagram of the system is as shown in Figure 4.

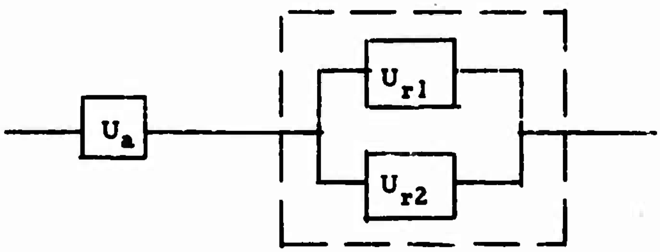


FIGURE 4
REDUNDANT CONFIGURATION - U_r
(Dissimilar Redundant Units)

U_a represents all units which are in series. The redundant configuration, U_r, is composed of two dissimilar redundant units, U_{r1} and U_{r2} which are equally effective in performing the required function. K_a will be used to designate the failure index of U_a, and K_{r1} and K_{r2} the failure indices of U_{r1} and U_{r2}, respectively. The method used to allocate reliability is to find a failure index for each redundant unit, K_r', so that for a given time period

$$R_r(K_r', K_r') = R_r(K_{r1}, K_{r2})$$

where $R_r(K_i, K_j)$ represents system reliability for a given time period, given redundant unit failure indices of K_i and K_j.

Given an equivalent failure index of K_r', equation (24) can be used to obtain K_r, the failure index of the redundant configuration, and the basic allocation equations then obtain. The following discussion is limited to redundant units which have approximately constant failure rates.

3.6.3.1 Active-Parallel Redundancy

Let $\bar{\lambda}$ represent the average failure rate of the normalizing function. The $K_j \bar{\lambda}$ represents the absolute failure rate of the j^{th} unit since K_j is obtained as the sum of component failure rates in the unit relative to the normalizing function. The reliability function for two units in an active-parallel redundant configuration is

$$R(t) = e^{-\lambda_{r1} t} + e^{-\lambda_{r2} t} - e^{-(\lambda_{r1} + \lambda_{r2})t}$$

where λ_{r1} and λ_{r2} are the failure rates of the redundant units. The problem then is to find a value of K_r' so that

$$e^{-K_{r1}\bar{\lambda}t} + e^{-K_{r2}\bar{\lambda}t} - e^{-(K_{r1} + K_{r2})\bar{\lambda}t} = 2e^{-K_r'\bar{\lambda}t} - e^{-2K_r'\bar{\lambda}t}$$

If we use the approximation $e^{-x} = 1 - x + x^2/2$, the above equation reduces to

$$K_r' = (K_{r1} \cdot K_{r2})^{1/2} \tag{25}$$

This approximate formula for K_r' is usually satisfactory. Equation (24) can then be used to obtain K_r and the allocations for the units in series, for the redundant units, and for the redundant configurations are obtained by the allocation equations given for serial or modified serial systems.

3.6.3.2 Standby Redundancy

From the general reliability function of a standby redundant configuration, K_r' must be determined so that

$$\frac{K_{r1}}{K_{r1} - K_{r2}} e^{-K_{r2}\bar{\lambda}t} - \frac{K_{r2}}{K_{r1} - K_{r2}} e^{-K_{r1}\bar{\lambda}t} = e^{-K_r'\bar{\lambda}t}(1 + K_r' \bar{\lambda} t)$$

The same approximation for e^{-x} as was used for active-parallel redundancy can also be employed to obtain an estimate for K_r'. This expression, however, will yield, for the right hand side, a term that involves $\bar{\lambda}^3 t^3$. Since $\bar{\lambda}$ will be quite small (say on the order of 20×10^{-6}), $\bar{\lambda}^3 t^3$ is negligible for the range of t usually involved. On dropping the term involving $\bar{\lambda}^3$, the approximate formula for K_r' is identical to equation (25). Tests of fifteen pairs of K_{r1} and K_{r2} showed an average error of 2% for K_r'. The maximum error was about 10% which occurred for the extremely unlikely ratio of $K_{r1}/K_{r2} = 100$.

Equation (24) can be used to obtain K_r and then basic allocation equations apply.

4.0 PROCEDURAL STEPS FOR ALLOCATION

4.1 Introduction

This section illustrates the step-by-step procedures for determining allocated unit reliabilities for: serial and modified serial systems, and redundant systems. The required data inputs are the relative functional failure rates and the modification factors reflecting the influences of special or nontypical conditions.

Two assumptions are required for proper utilization of the allocation methods. They are:

(1) That allocation levels can be so chosen that failure probabilities are independent, i.e., dependent components can be grouped into one unit, making failure probability of this unit independent of the state of the other units.

(2) That the unit state can be described in discrete terms of success and failure through analysis of the system reliability requirements and the functional relationship between unit and system operation.

The procedural steps have been designed to accommodate each of the specific system types. Attention is called to the fact, however, that five steps are applicable to all systems. These are described in Section 4.2, and are preliminary to any steps that are peculiar to the specific system types. Steps applying to serial and modified serial systems will be found in Section 4.3. Steps for redundant systems are described in Section 4.4.

4.2 Initial Steps (Applicable to all Systems Types)

The following steps apply to all of the system types considered herein and must be accomplished first. The steps peculiar to each system type are to be performed in the order listed. (See Sections 4.3 and 4.4.) The sample allocation work sheet, Figure 5, can be used for all system types.

(a) Define the units for which the system reliability requirement is to be allocated by constructing a reliability block diagram showing the units (blocks) of the system in logical sequence.

(b) Determine the system type (serial or modified serial, redundant or bimodal) by referring to the definitions given in Section 2.4. The term redundant configuration shall apply to the group of redundant units.

(c) Obtain unit failure indices K_j. These K_j's should be listed on an allocation worksheet similar to that of Figure 5.

(d) From the definition of system success and the operational demands imposed on each unit, estimate the essentiality, E_j, of each series unit and the redundant configuration if applicable. (For unmodified serial systems, E is equal to one for all units.) List the unit essentialities in the appropriate column of the worksheet.

DESIGN STAGE RELIABILITY ALLOCATION WORKSHEET

System ______________________ Date ______________

Primary Mission ________________ System Operating Time - T ______________

System Reliability Requirement † ______________

		Series Units				Redund. Config.	Redundant Units	
Unit or Configuration Identification	U_j	U_1	U_2	. . .	U_m	U_r	U_{r1}	U_{r2}
Essentiality	E_j	E_1	E_2	. . .	E_m	E_r		
Operating Time	t_j	t_1	t_2	. . .	t_m	t_r		
Failure Index	K_j	K_1	K_2	. . .	K_m	K_r †	K_{r1}	K_{r2}
Failure Index Ratio	w_j †	w_1	w_2	. . .	w_m	w_r	w_{r1}	w_{r2}
Allocated Reliability	$\hat{R}(t_j)$ †	$\hat{R}(t_1)$	$\hat{R}(t_2)$	. . .	$\hat{R}(t_m)$	$\hat{R}(t_r)$	$\hat{R}_{r1}(t_r)$	$\hat{R}_{r2}(t_r)$

† Obtained according to appropriate steps of procedures

FIGURE 5
SAMPLE ALLOCATION WORKSHEET

(e) If the system requirement is in terms of a probability of successful operation for T hours, estimate the average operating time, t_j, of each unit and the redundant configuration during T system hours of operation. If the requirement is in terms of mean life or failure rate, a value of T should be chosen to represent a significant period of system operation such as average mission-time or average maintenance-period. Provision for listing the t_j's is also made on the worksheet.

After the completion of the above steps, refer to Sections 4.3 or 4.4 for the specific system type under consideration.

4.3 Final Steps (Serial or Modified Serial Systems)

Assuming that the preliminary steps described in Section 4.2 have been completed, then the following steps for serial or modified serial systems should be performed in the order listed.

The identification notation shown in Figure 6 will be used,

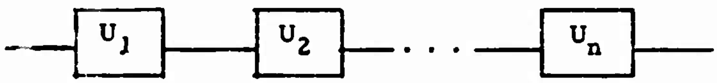


FIGURE 6

NOTATION CONVENTION

where U_1 to U_n are series or modified series units.

1. Determine the system reliability requirement by first estimating system design capability. If the design capability is less than one (1.0), the original system requirement shall be considered to be a system effectiveness requirement unless otherwise specified. The equivalent system reliability

requirement is given by 6/

$$R*(T) = \frac{S*(T)}{D_s} \qquad (26)$$

where $R*(T)$ is the system reliability requirement for T system or mission hours of operation.

$S*(T)$ is the system effectiveness requirement for T system or mission hours of operation.

D_s is the system design capability.

NOTE: If the original requirement is given in terms of mean life or failure rate, a value of T should be chosen to represent a significant period of system operation such as average mission time or average maintenance period. The equivalent system effectiveness requirement is obtained from the equation

$$S*(T) = e^{-T/\bar{\theta}*} \qquad (27)$$

or

$$S*(T) = e^{-\lambda*T} \qquad (28)$$

where

$\bar{\theta}*$ is the original system mean life requirement.

$\lambda*$ is the original system failure rate requirement. Equation (26) is then used to determine the system reliability requirement.

2. Determine the feasibility of the reliability requirement through consideration of past reliability performance. Procedures based on normalized complexity and environmental measures and reliability prediction by function are appropriate.

6/ This equation is based on the assumption that operational readiness or availability is equal to one. Since availability is a function of reliability as well as maintenance, one must consider such factors as the mission profile, the ratio of repair rates to failure rates, and the availability of backup equipment in interpreting $R*(T)$.

3. Obtain the total system failure index, K, by the formula

$$K = K_1 + K_2 + \cdots + K_j + \cdots + K_n \tag{29}$$

and form the failure index ratios

$$\begin{aligned} w_1 &= K_1/K \\ w_2 &= K_2/K \\ &\cdot \\ &\cdot \\ &\cdot \\ w_n &= K_n/K \end{aligned} \tag{30}$$

4. Compute allocated unit reliabilities from the equation

$$\hat{R}(t_j) = 1 - \frac{1 - R^*(T)^{w_j}}{E_j} \tag{31}$$

where

$\hat{R}(t_j)$ is the allocated reliability of the j^{th} unit.

t_j is the average operating time of the j^{th} unit during T hours of system operation.

E_j is the essentiality of the j^{th} unit. (E_j must be greater than $1 - R^*(T)^{w_j}$. Units which violate this requirement shall be excluded from the allocation and new values of the w's obtained.)

5. If unit requirements are desired in terms of mean life or failure rate and constant failure rates are assumed,

$$\hat{\theta}_j = - \frac{t_j}{\log \hat{R}(t_j)} \qquad \text{(natural logarithms)} \tag{32}$$

$$\hat{\lambda}_j = - \frac{\log \hat{R}(t_j)}{t_j} \tag{33}$$

Approximate Formulas:

Follow Steps 1 to 3;

For A. Unit mean life requirements:

$$\hat{\theta}_j = -\frac{E_j t_j}{w_j \log R^*(T)} \tag{34}$$

For B. Unit failure rate requirements:

$$\hat{\lambda}_j = -\frac{w_j \log R^*(T)}{E_j t_j} \tag{35}$$

For units known to have a failure rate which is not constant over time, an average failure rate during the unit's t hours of operation can be obtained by the equation

$$\hat{\lambda}_j = \frac{1-\hat{R}(t_j)}{t_j} \tag{36}$$

4.4 Final Steps (Redundant Systems)

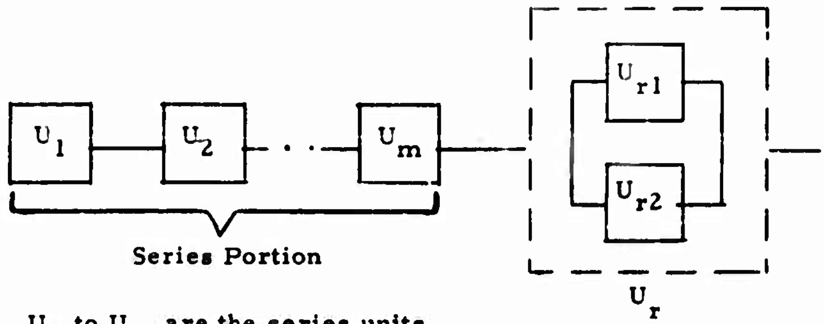
This section describes the remaining steps of the allocation procedure for systems which contain a single redundant configuration consisting of two units, not necessarily identical, each of which is equally effective in performing the required function--i.e., modal design capabilities are equal. The two types of redundancy considered are:

(a) Active-parallel, where both redundant units are constantly energized, and

(b) Standby, where one redundant unit is not energized until the operating redundant unit fails.

The method used for allocating reliability in these cases is to determine an equivalent failure index for the redundant configuration from the failure indices of each redundant unit. This procedure will permit utilization of the basic allocation formulas. If switching is involved, it shall be assumed that the probability of premature switching (switching when not required) is relatively small as compared to the probability of failure to switch when required. The switching mechanism, therefore, can be considered as a series unit.

The unit identification notation shown in Figure 7 will be used in this section.



U_1 to U_m are the series units

U_{r1} and U_{r2} are the two redundant units

U_r is the redundant configuration

FIGURE 7

NOTATION CONVENTION OF SECTION 4.4

The remaining steps of the procedure appear in sequence below:

1. Determine the system reliability requirement by first estimating the design capability of the system. (Follow Step 1 of Section 4.3 except that system design adequacy, D_s, is taken to be the design capability of each of the two modes of operation.)

2. Obtain the total failure index of all series units from the equation

$$K_a = \sum_{j=1}^{m} K_j \tag{37}$$

where

K_a is the total failure index of the series portion

m is the number of series units.

3. If the redundant units are not identical, calculate an equivalent failure index for each redundant unit from the formula

$$K'_r = \left[K_{r1} \cdot K_{r2} \right]^{1/2} \tag{38}$$

where

K'_r is the equivalent failure index of redundant units having failure indices of K_{r1} and K_{r2}. (If U_{r1} and U_{r2} are identical, K'_r is the failure index for each of them.)

4. Determine the feasibility of the reliability requirement.

5. Calculate the ratio

$$\alpha = \frac{K_a}{K'_r} \tag{39}$$

6. To obtain an equivalent series failure index for the redundant configuration,

A. <u>With Active Parallel Redundancy:</u>

Obtain the value $Z(\alpha, R^*)$ from Figure 8 or 9 for the appropriate set of α and R^*,

B. <u>With Standby Redundancy:</u>

Obtain the value $Z(\alpha, R^*)$ from Figure 10 or 11 for the appropriate set of α and R^*, and compute

$$K_r = Z(\alpha, R^*) K_a \tag{40}$$

where K_r is the equivalent series failure index of the redundant configuration.

7. Obtain the total system failure index from the equation

$$K = K_a + K_r \tag{41}$$

and calculate the series unit and redundant configuration failure index ratios:

$$w_1 = K_1/K$$
$$w_2 = K_2/K$$
$$\vdots \qquad (42)$$
$$w_m = K_m/K$$
$$w_r = K_r/K$$

as well as each redundant unit failure index ratio:

$$w_{r1} = K_{r1}/K$$
$$w_{r2} = K_{r2}/K \qquad (43)$$

8. Allocated reliabilities for T system hours of operation can be computed from the equations given below.

A. <u>Series Units</u>

The allocated reliability for the j^{th} series unit is

$$\hat{R}(t_j) = 1 - \frac{1-R^*(T)^{w_j}}{E_j} \qquad (j = 1, 2, \ldots, m) \qquad (44)$$

B. <u>Redundant Configuration</u>

The allocated reliability for the redundant configuration is

$$\hat{R}(t_r) = 1 - \frac{1-R^*(T)^{w_r}}{E_r} \qquad (45)$$

where

t_r is the average operating time of the redundant configuration during T hours of system operation.

E_r is the essentiality of the redundant configuration. (E_r must be greater than $1-R^*(T)^{w_r}$.)

C. Redundant Units

The allocated reliability for each redundant unit is

$$\hat{R}_{ri}(t_r) = R^*(T)^{w_{ri}} \qquad (i = 1, 2) \tag{46}$$

▷ 9. If unit requirements are desired in terms of mean life or failure rate, Step 5 of Section 4.3 applies for the series and redundant units, provided the assumption of constant unit failure rate is good. If a constant failure rate is a poor assumption, average unit or redundant configuration failure rate can be allocated by the equation

$$\hat{\lambda}_j = \frac{1 - \hat{R}(t_j)}{t_j} \tag{47}$$

For the redundant configuration, mean life requirements are computed from the following equations:

Active Parallel

$$\hat{\theta}_r = \hat{\theta}_{r1} + \hat{\theta}_{r2} - \frac{\hat{\theta}_{r1}\,\hat{\theta}_{r2}}{\hat{\theta}_{r1} + \hat{\theta}_{r2}} \tag{48}$$

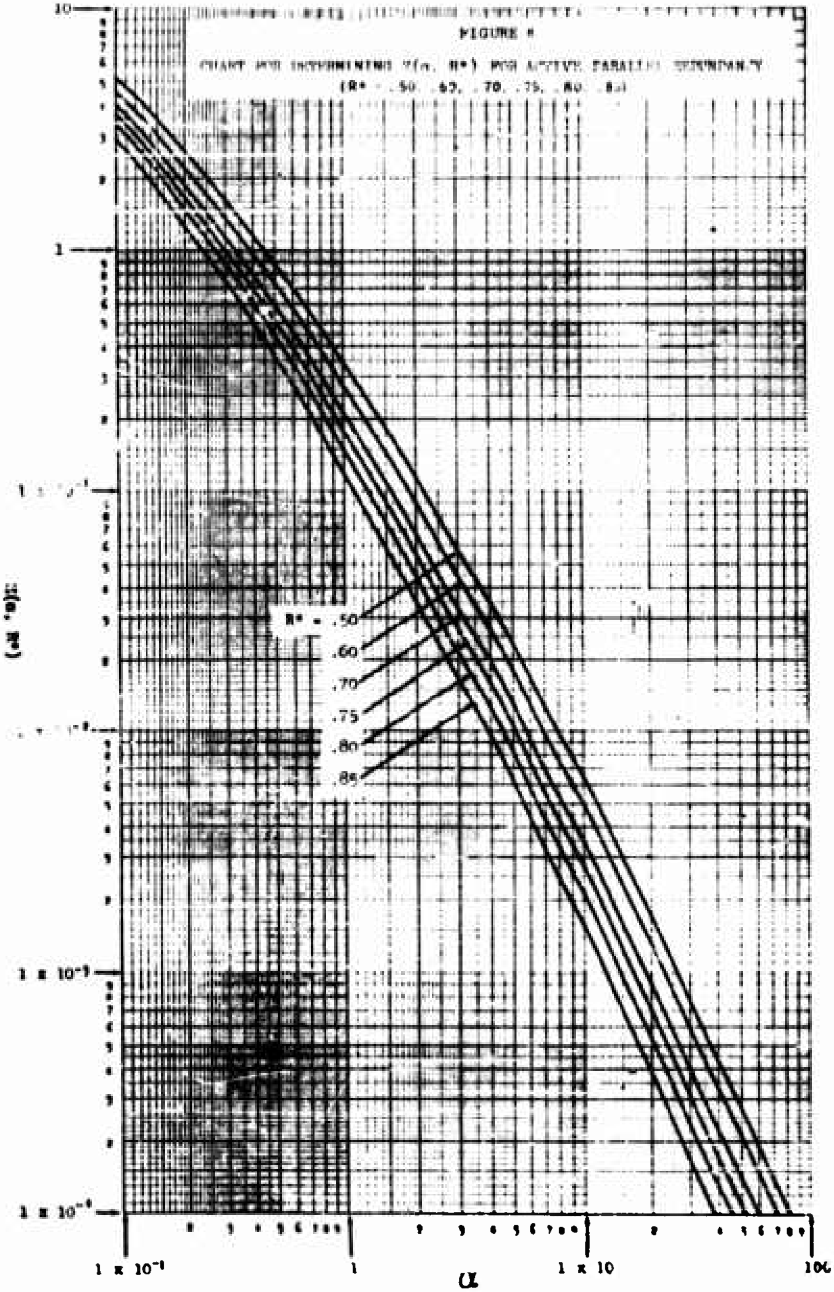
Standby

$$\hat{\theta}_r = \hat{\theta}_{r1} + \hat{\theta}_{r2} \tag{49}$$

NOTE: The time-to-failure distribution of a redundant configuration is not truly exponential. $\hat{\theta}_r$, therefore, cannot be interpreted as the mean life associated with a constant failure rate.

The average failure rate of the redundant configuration is given by

$$\hat{\lambda}_r = \frac{1 - \hat{R}(t_r)}{t_r} \tag{50}$$



R* = .50
.60
.70
.75
.80
.85
10
1
1 x 10⁻³
1 x 10⁻⁴
1 x 10⁻¹
1
1 x 10
α

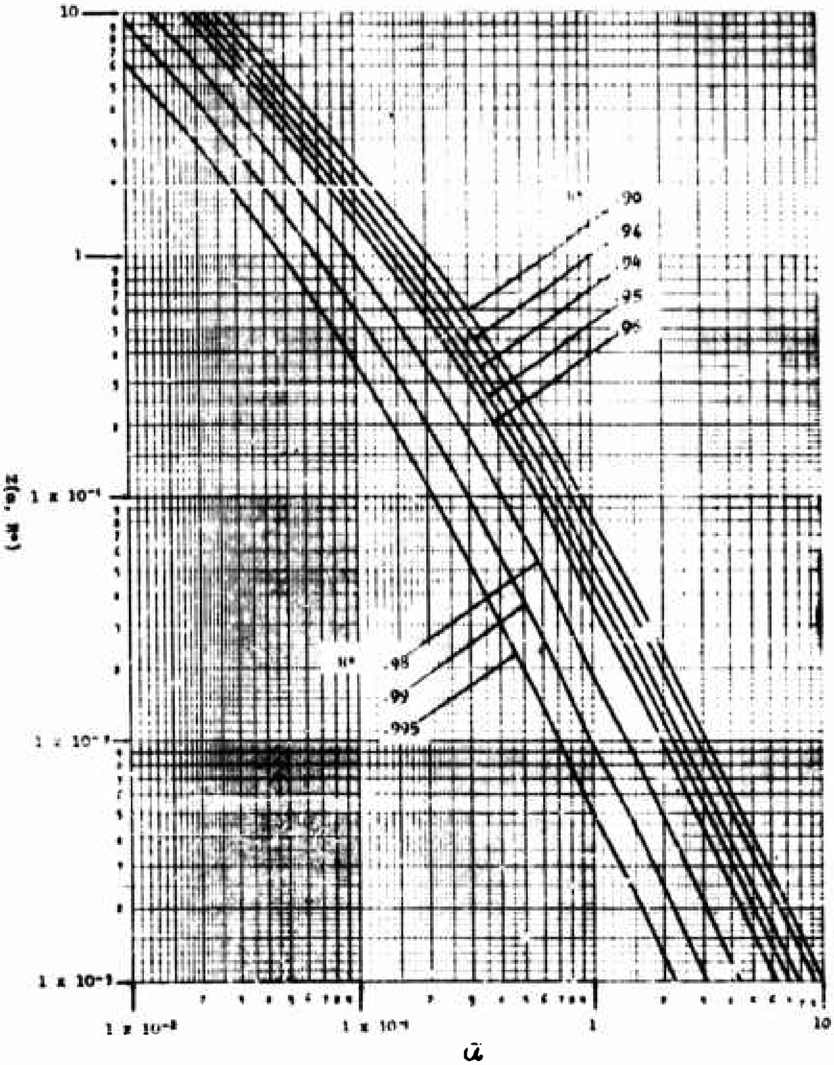


FIGURE 9

CHART FOR DETERMINING Z(α, R*)
FOR ACTIVE PARALLEL REDUNDANCY
(R* = .90, .92, .94, .95, .96, .98, .99, .995)

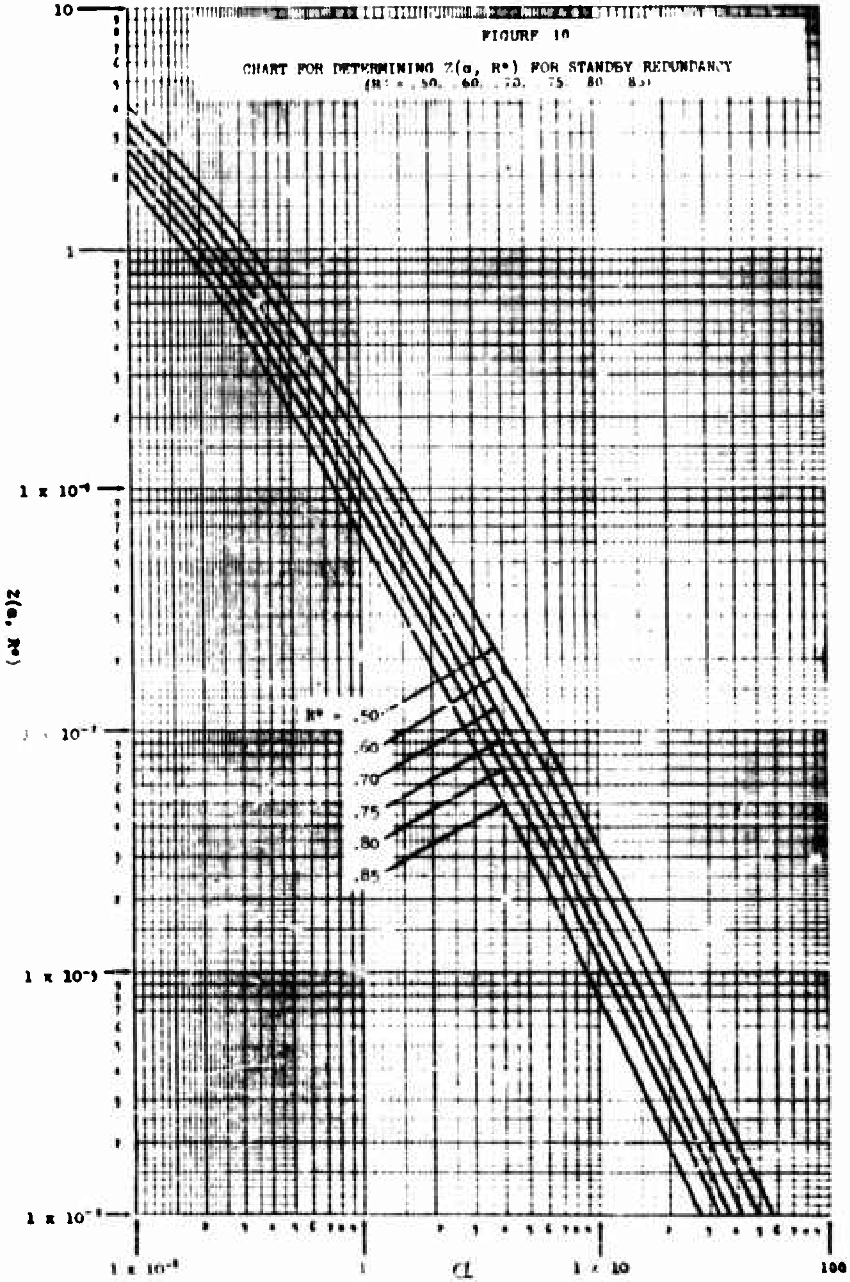


FIGURE 10
CHART FOR DETERMINING Z(a, R*) FOR STANDBY REDUNDANCY
R* = .50
.60
.70
.75
.80
.85
Z(a, R*)
a
100

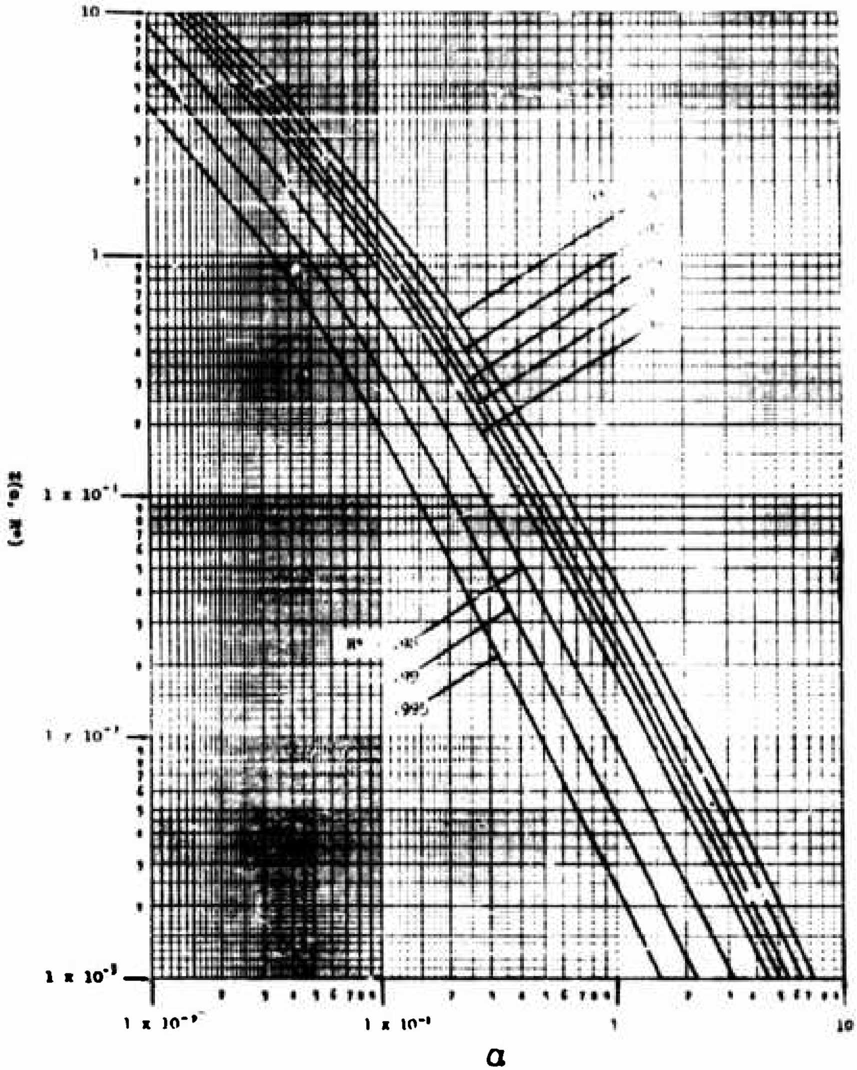


FIGURE 11

CHART FOR DETERMINING Z(α, R*)
FOR STANDBY REDUNDANCY
(R* = .90, .92, .94, .95, .96, .98, .99, .995)

EXAMPLE C

BALLISTIC MISSILE PAYLOAD ALLOCATION[1/]

1/ CONDENSED FROM "A PRELIMINARY DESIGN AID FOR STUDYING COMPONENT WEIGHT ASSIGNMENTS IN BALLISTIC MISSILE PAYLOADS," By S. I. Firstman, RAND Memorandum RM-2471, January 13, 1960.

ABSTRACT

Each element of a ballistic missile's payload--warhead, guidance and penetration aids--will increase in effectiveness with an increase of weight allocated to the element. For a missile that is to be employed against a defended "point" target, this example presents a method for determining the optimum division of the missile's payload between the three competing (for weight) elements, when their individual weight-effectiveness relationships are known. For the case of a single missile per target, using a most basic application of the stepwise optimization philosophy of dynamic programming, the problem is formulated as a two-stage weight allocation process. The first stage determines the optimum tradeoff between warhead (lethal radius) and guidance (CEP); the second stage determines the optimum division between penetration aids and an optimum mix of warhead and guidance. The simple arithmetical method that results is demonstrated by an example. The same optimization process is useful for the cases of sequential and simultaneous multiple missile employment per target. Although this design optimization problem can be solved, functionally, for the modes of missile employment considered, its applicability to a real allocation problem is confounded by the design, intelligence and employment estimates required in the analysis. Use of this method could show, however, the influence of the estimate uncertainties on the optimal payload division and could thereby serve as a useful point of departure for design compromises.

TABLE OF CONTENTS

ILLUSTRATIONS

TABLES

1.0 INTRODUCTION

When determining the design parameters of an item of equipment, it is often desirable to employ a quantitative model that describes or predicts the equipment's capability in terms of the relevant parameters. This model, though sometimes relatively crude, affords a means of determining the optimum, or nearly so, set of design parameters. Ballistic missile payloads are a case in point, where one convenient model of capability is the missile's potential ability to survive enemy defenses and damage or destroy what is called, a hardened "point" target; given that it is delivered to the target area in a nonfailed condition. For this model of capability, the missile payload design-parameter-optimization process is a simple numerical procedure. It is developed and demonstrated in this example.

Each element of a ballistic missile's payload--guidance, warhead, and penetration aids--will increase in capability with an increase of weight allocated to the element. The ability to destroy a "point" target is dependent on the ability of the missile to impact within the lethal radius of the target. This destruction capability, therefore, is dependent upon: (a) the guidance accuracy, which can be defined as a function of the guidance system weight, and (b) the target lethal radius, which for a fixed target hardness can be defined as a function of the missile warhead yield, which in turn is dependent upon the warhead weight. 2/

The ability to survive the enemy defenses is dependent upon: (a) the offensive tactic employed, (b) the types, characteristics, and numbers of the penetration aids, (c) the type of defense, its strength, and its ability to cope with the penetration aids. To determine the probability of surviving enemy defenses as a function of these several variables is indeed a difficult

2/ The general relationship of these factors to availability and dependability is illustrated in Volume III, Example B, TG-II Report.

task and is presently further aggravated by many technical and operational uncertainties. However, persons studying this penetration problem feel that, to a first-order approximation, the ability to survive ICBM defenses can be described as a function of the weight devoted to penetration aids.

Starting with the weight-effectiveness relationships for each of the competing (for weight) elements, the problems of determining the optimum division of payload for both single and multiple (sequential and simultaneous) missile employment per target will be formulated and solved using a most basic application of the stepwise optimization philosophy of dynamic programming. The simple arithmetical method that results will then be demonstrated by an example. Following that, the uncertainties surrounding the true operational context and the difficulties of making precise pre-design performance estimates will be considered to indicate more clearly the limitations on the utility of the method developed.

2.0 SINGLE MISSILE PER TARGET

2.1 Problem Formulation

A fixed missile payload, W, is to be divided among three systems, guidance, warhead, and penetration aids. The weight allocated to each system must, for physical (and operational) reasons, satisfy some minimum requirement,

guidance,	$w_g \geq w_{g_o}$
warhead,	$w_w \geq w_{w_o}$
penetration aids,	$w_p \geq w_{p_o}$

and be at levels such that the total payload is

$$W = w_g + w_w + w_p \tag{1}$$

The intent of the allocation is to maximize the missile's potential offensive capability, which is defined as the probability that a missile destroys a particular defended point target; given that it is delivered to the target area in a nonfailed condition.[3/] Neglecting reliability considerations, since it is assumed that each element will be made as reliable as possible for a given weight,[4/] this measure of effectiveness is given by

$$P = p_c \, p_k \qquad (2)$$

where:

$p_s \equiv P$ (the missile survives enemy defenses).

$p_k \equiv P$ (the missile falls within the target lethal radius), i.e., the single-shot kill probability.

These two probabilities are independent, and both are functions of their weight allocations; p_s is a monotonically increasing function of w_p; and p_k is a nonlinear function of w_g and w_w. The payload division problem shall be formulated and solved using a two-step dynamic programming stepwise optimization technique that for this problem is simply a directed search over combinations of allocations.

2.2 Method of Solution

The first stage in the allocation process is to examine the tradeoff between guidance accuracy and warhead yield and determine the levels of w_g and w_w which, for each fixed weight assignment will maximize p_k. For a circular normal impact distribution and assuming a "cookie-cutter" destruction distribution,[5/] p_k is given by

---

3/ In general, the effectiveness of each missile of the type being designed is to be maximized with respect to the characteristics of a particular class of targets.

4/ Depending on the use made of this design aid, the weight estimate employed in the analysis should either be sufficiently gross so as to allow for minor changes in design for reliability improvement purposes (preliminary design of new system), or sufficiently precise that no changes in equipment are likely (marriage of off-the-shelf items).

5/ The "cookie-cutter" destruction distribution assumes a dichotomy of lethality due to blast damage from a nuclear weapon; targets of a given hardness that lie within the lethal radius of the weapon are destroyed while targets outside the lethal radius (or "cookie-cutter") are not even damaged.

$$p_k = 1 - 2^{-(LR/CEP)^2} \tag{3}$$

where

LR is the lethal radius of the target hardness-missile yield combination, and

CEP is the circular error probable of the impact distribution (equal to 1.177 times the standard deviation of miss distance).

For each level of $W \geq w_{g_o} + w_{w_o}$, let

$$p_k(W) = \max_{\substack{w_g \geq w_{g_o} \\ w_w \geq w_{w_o}}} \left[1 - 2^{-(LR/CEP)^2}\right] \tag{4}$$

where

$$W = w_g + w_w \tag{5}$$

Due to the form of the function, the problem of finding that combination of w_g and w_w that maximizes $p_k(W)$ can be seen to be the same as finding the maximum ratio LR/CEP for the given W.

By letting the functions defining the LR and CEP be $LR = h(w_w)$, and $CEP = g(w_g)$, the problem becomes: find those levels of w_w and w_g that maximize

$$f(w) = \frac{h(w_w)}{g(w_g)} \tag{6}$$

subject to

$$w_g \geq w_{g_o}$$

$$w_w \geq w_{w_o}, \text{ and}$$

$$w_w + w_g = W$$

Then, for the maximum level of $f(W)$,

$$p_k(W) = 1 - 2^{-\left[f(W) \max\right]^2} \tag{7}$$

If $h(w_w)$ and $g(w_g)$ were well behaved and differentiable throughout their range, then analytical methods could be employed for this problem. This, however, need not be the case, as these dependencies could be described by step functions, or indeed may be just several discrete values representing several existing designs.

For discrete levels[6] of w_g and w_w, because of the form of f(W), this allocation problem can be readily solved numerically using a simple and fairly rapid search over the range of combinations of w_w and w_g possible for each W. Formally, this search process is a basic application of Bellman's[7] method of examining a series of successive approximations in policy space. This method will be demonstrated by an example.

Having obtained $p_k(W)$ for several levels of W, this information can be utilized to find that level of w_p which will give

$$P(W) = \max_{w_{p_o} \leq w_p \leq W} \left[p_s(w_p)\, p_k(W - w_p) \right] \tag{8}$$

This second-stage allocation problem can be solved by examining the range of possible allocations to w_p and an optimal combination of w_g and w_w.[8] For each level of $(W - w_p)$, the combination that yields the maximum p_k is known from the first-stage of the problem and therefore the combination of $(W - w_p)$ and w_p that yields a maximum product, for each level of W, is the optimum combination. The optimization method, which is similar in nature to that employed in the first stage, will also be demonstrated in the example.

6/ A discrete approximation is employed if the functions are continuous.

7/ Op. Cit. Bellman, Richard, "Dynamic Programming," Princeton University Press, 1957.

8/ If more than one defense mode is anticipated, $p_s(w_p)$ could be the result of an appropriate sub-optimization process.

3.0 MULTIPLE MISSILES PER TARGET

3.1 Simultaneous Delivery

The preceding analysis was based upon the use of a single warhead per missile and a single missile per target. If multiple missiles of identical design, each with a single warhead, are employed simultaneously against a target, it appears reasonable to employ as an objective function that is to be maximized

$$P_n \equiv P\left[\begin{array}{l}\text{at least one of n missiles survives and destroys the} \\ \text{target .}\end{array}\right]$$

Assuming non-correlated impact errors and non-additive destruction effects, this can be written as

$$P_n = 1 - (1 - p_s^{(n)} p_k)^n \tag{9}$$

where,

$$p_s^{(n)} \equiv P\left\{\begin{array}{l}\text{Survival of each missile when n are simultaneously} \\ \text{employed.}\end{array}\right]$$

By inspection it can be seen that P_n will be a maximum when $p_s^{(n)} p_k$ is a maximum. The levels of w_g, w_w, and w_p that maximize $p_s^{(n)} p_k$ can be obtained as before, when $p_s^{(n)}$ is known.

No restrictions are necessary on the form of $p_s^{(n)}$ for this analysis, but it should be noted that if multiple missiles are employed simultaneously, they should add mutual support to each other in penetrating the enemy defenses. It appears plausible to expect that since the effectiveness of penetration aids can be expressed in terms of pounds of aids employed for a single missile, the same type of relationship can be defined for multiple missile employment. Where the preceding analysis implicitly employed one curve describing p_s as a function of w_p, multiple warhead employment would lead to a family of curves similar to those indicated in Figure 1.

For this case, then, depending on the anticipated employment, several sets of optimum allocations could be obtained for each payload weight. In order to be of use in the design process, an analysis using the method probably would need to be done when the missile is in the preliminary design stage. It does not appear likely that the number of missiles

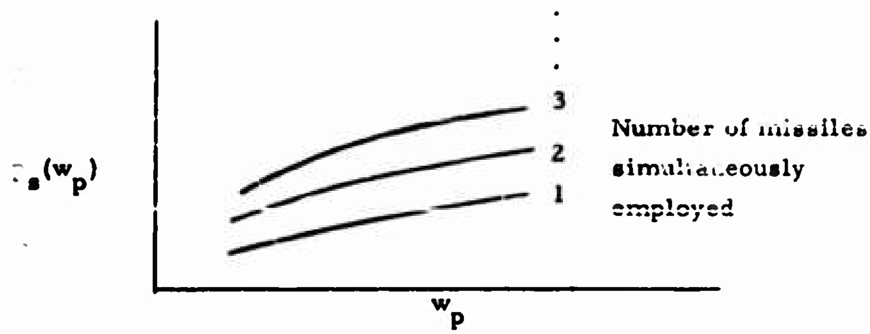


FIGURE 1. MULTIPLE MISSILE SURVIVAL

that will be employed against a particular target would be known at that time. Indeed, even the number of such missiles, to be procured and emplaced probably would not be known at that time. Moreover, because of failures during launch or powered flight, or because of enemy action, the number of missiles that is actually delivered simultaneou ly to the target area may be different than the number planned. Therefore, a compromise based perhaps on some plausible or conservative number of missiles per target probably would be necessary.

In considering the simultaneous employment of purely penetration-aid missiles (no warhead) and purely warhead missiles (no penetration aids), the form of the objective function employed above would need to be modified to

$$P_n = 1 - (1 - p_s^{(m,n)} p_k)^n \qquad (10)$$

where,

$p_s^{(m,n)} \equiv P$ [survival of each warhead-carrying missile when m penetration aid and n warhead missiles are simultaneously employed.]

As before, P_n will be a maximum when $p_s^{(m,n)} p_k$ is a maximum. Under the assumptions used, the design of the warhead missile will be optimum at the levels of w_g and w_w that maximize p_k; and this can be obtained as before. On the other hand, $p_s^{(m,n)}$, in addition to depending on the levels of m and n, would be a function of the mix between penetration aids and guidance on the pentration-aid missile.

3.2 Sequential Delivery

Multiple missiles can also be employed in a sequential manner against a target. In this case, because of maintenance (a particular missile may be "down" awaiting maintenance when hostilities begin), and the operational and reliability considerations previously alluded to, it does not appear plausible to assign a rigid a priori sequence to a set of missiles that are to be directed against a particular target. A fixed sequence could be difficult to obtain operationally. Therefore, this analysis will be based upon the assumption that all missiles of a class will have the same design parameters rather than special payload designs geared to the anticipated sequence of employment. This argument is strengthened by the consideration that because of the changing pattern of targets and of weapon demands, the number of weapons to be programmed against a target is probably also time-variant. With these considerations in mind, then, the analysis will be directed to find an optimum division of payload that is independent of sequence of launch and of the number launched.

Considering the first case where two missiles are employed, and changing notation slightly, for the first missile,

$$P_1 \equiv P\left[\text{first missile survives the defenses and destroys the target}\right]$$

which is, as before

$$P_1 = p_s^{(1)} p_k \tag{11}$$

where

$$p_s^{(1)} \equiv P\left[\text{first missile survives}\right]$$

For the second missile, assuming no additive effects of destruction so that all the p_k are identical

$$P_2 = p_s^{(2)} p_k \tag{12}$$

where, by decomposition

$$p_s^{(2)} = p_s^{(2/1)} p_s^{(1)} + p_s^{(2/\bar{1})}(1 - p_s^{(1)}) \tag{13}$$

where

$$p_s^{(2/1)} \equiv P\left[\text{missile two survives given that missile one survived}\right]$$

$$p_s^{(2/\bar{1})} \equiv P\left[\begin{array}{l}\text{missile two survives given that missile one did not}\\ \text{survive}\end{array}\right]$$

Therefore,

$$P_2 = \left\{p_s^{(2/1)} p_s^{(1)} + p_s^{(2/\bar{1})}(1 - p_s^{(1)})\right\} p_k \tag{14}$$

By making the conservative assumption that the enemy's missile defense has no weaknesses, e.g., has no rate-of-fire or stockpile limitations,[9] it can be stated that

$$p_s^{(2/\bar{1})} = p_s^{(1)} \tag{15}$$

Then

$$\begin{aligned} P_2 &= p_k \left\{p_s^{(2/1)} p_s^{(1)} + p_s^{(1)} - (p_s^{(1)})^2\right\} \\ &= P_1 \left\{p_s^{(2/1)} + 1 - p_s^{(1)}\right\} \\ &= P_1 + P_1 \left\{p_s^{(2/1)} - p_s^{(1)}\right\} \end{aligned} \tag{16}$$

where it appears reasonable to assume that

$$p_s^{(2/1)} \geq p_s^{(1)}, \tag{17}$$

and following from the previous assumptions about the enemy defenses,

$$p_s^{(2/1)} > p_s^{(1)} , \tag{18}$$

only if the first missile damaged the defenses.

---

[9] If it is postulated that the enemy's defenses would have either rate-of-fire or stockpile limitations, the sequential employment of penetration aid-carrying missiles followed by warhead-carrying missiles appears to be interesting. However, the desirability of that tactic and the division of the penetration-aid missile payloads are problems beyond the scope of this example. Under the mode-of-destruction assumptions employed, the warhead missile's payload would obviously be designed for maximum p_k as before.

Let

$$\Delta p_s = p_s^{(2/1)} - p_s^{(1)}; \quad 0 \leq \Delta p_s \leq 1 \tag{19}$$

then

$$P_2 = P_1 + P_1 \Delta p_s \tag{20}$$

where it can be reasoned that Δp_s is determined primarily by the enemy defenses.

For two missiles, employing the same destruction assumptions as before, it appears that a reasonable objective is to maximize

$$\begin{aligned} P &= P\left[\text{at least one missile survives the defenses and destroys the target}\right] \\ &= 1 - (1 - P_1)(1 - P_2) \\ &= 1 - (1 - P_1)(1 - P_1 - P_1 \Delta p_s) \\ &= 2P_1 - P_1^2 + P_1 \Delta p_s - P_1^2 \Delta p_s \end{aligned} \tag{21}$$

This means that the over-all probability of mission success is dependent on both P_1 and Δp_s. But, Δp_s is dependent primarily on the defenses (how they are built, operated, etc.), and therefore, the offense should probably plan on the worst case, which is $\Delta p_s = 0$. This means that the defenses are totally unaffected by the employment of the first weapon.

Employing this conservative operational assumption then, the problem becomes that of choosing levels of w_g, w_w, and w_p so as to maximize

$$P = 2P_1 - P_1^2$$

This is seen to be the probability that either of two missiles destroys the target, if each missile is of the same design and must penetrate the same defenses. This function increases monotonically with P_1; is a maximum when P_1 is a maximum, and therefore, the single missile per target data and optimization method are applicable to this situation. Although developed for the two-weapon case, it can be seen by induction that this result is applicable to all numbers of sequential missiles as long as the conservative assumptions relative to effects on defenses and destruction phenomena remain reasonable.

4.0 AN EXAMPLE

Assume that for a defended point target of given hardness, the functions $g(w_g)$, $h(w_w)$ and $p_s(w_p)$ are as given in Figure 2. The first step in the weight allocation process is to find $p_k(W)$, for several levels of W. This is done in Table I.

TABLE I

DETERMINATION OF MAXIMUM p_k

W (lb)	f(W)max	Optimum Sub-Allocation		$[f(W)max]^2$	$2-[f(W)max]^2$	$p_k(W)$
		w_w	w_g			
500	0.22	200	300	0.048	0.97	0.03
600	0.25	200	400	0.063	0.96	0.04
700	0.27	200	500	0.073	0.95	0.05
800	0.32	200	600	0.102	0.93	0.07
900	0.34	300	600	0.116	0.92	0.08
1000	0.38	300	700	0.145	0.90	0.10
1100	0.42	300	800	0.176	0.88	0.12
1200	0.47	200	1000	0.221	0.86	0.14
1300	0.54	200	1100	0.292	0.82	0.18
1400	0.62	200	1200	0.385	0.77	0.23

For example f(W) for W = 500 is fixed by the arbitrary constraints on w_g and w_w, w_{g_o} = 300 lb and w_{w_o} = 200 lbs. hence,

$$f(W)\Big|_{W = 500} = \frac{h(200)}{g(300)} = \frac{0.43}{2.00} = f(W)max$$

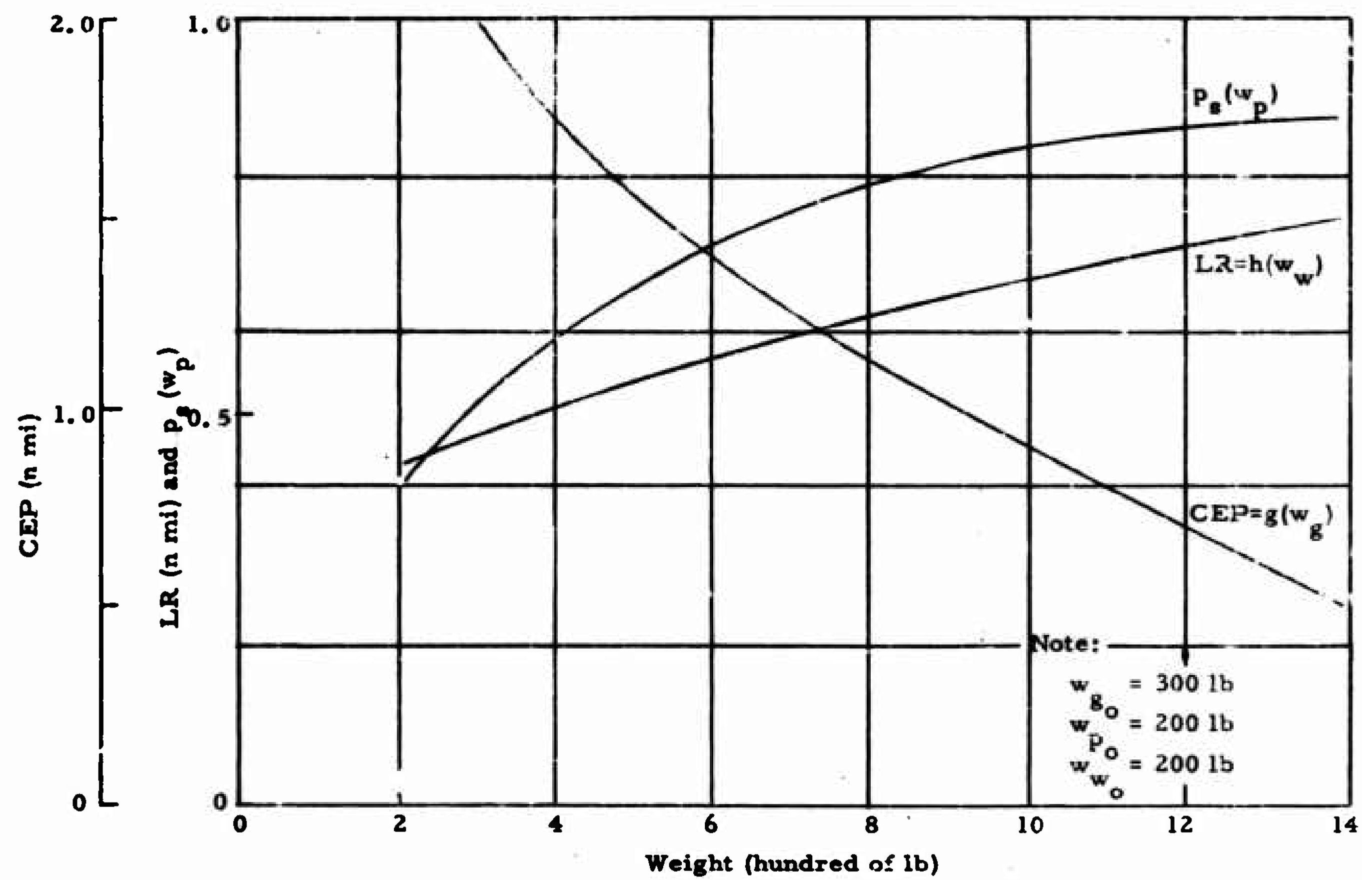


FIGURE 2

CEP, LR, AND p_s AS A FUNCTION OF WEIGHT ALLOCATION

$f(W)\big|_{W=600}$ is the maximum of the two combinations

$$\frac{h(200)}{g(400)} = \frac{0.43}{1.75};\ \frac{h(300)}{g(300)} = \frac{0.48}{2.00}$$

$f(W)\big|_{W=700}$ is the maximum of the three combinations

$$\frac{h(200)}{g(500)} = \frac{0.43}{1.57};\ \frac{h(300)}{g(400)} = \frac{0.48}{1.75};\ \frac{h(400)}{g(300)} = \frac{0.52}{2.00}$$

As can be seen, this process is straightforward and quite rapid.

The value of $p_k(W)$ as a function of W is now known. The second step uses this maximum p_k and the associated mix between w_g and w_w to obtain the maximum value of P, for each level of W, or a particular value of W. The procedure for obtaining P(W) is shown in Table II (reference equation 8).

TABLE II

DETERMINATION OF MAXIMUM P

W (lb)	w_w [1]	w_g [1]	$p_k(W)$ [1]	P(W)	w_p
500	200	300	0.03	--	--
600	200	400	0.04	--	--
700	200	500	0.05	0.01	200
800	200	600	0.07	0.02	200
900	300	600	0.08	0.02	300
1000	300	700	0.10	0.03	200 [2]
1100	300	800	0.12	0.04	300
1200	200	1000	0.14	0.04	400
1300	200	1100	0.18	0.05	300
1400	200	1200	0.23	0.06	300
1500				0.07	300

[1] Repeated from Table I.

[2] This anomaly is caused by the jump of $p_k(W)$ from 0.05 to 0.07, which in turn is a result of the number of significant figures employed.

For example $P(W)|_{w=700}$ is fixed by the arbitrary constraints on w_w, w_g and w_p.

$$P(W)|_{w=700} = \left[p_s(200)\right]\left[p_k(500)\right] = (0.40)(0.03)$$

$P(W)|_{w=800}$ is the maximum of the two combinations

$$\left[p_s(200)\right]\left[p_k(600)\right] = (0.40)(0.04)$$
$$\left[p_s(300)\right]\left[p_k(500)\right] = (0.52)(0.03)$$

$P(W)|_{w=900}$ is the maximum of the three combinations

$$\left[p_s(200)\right]\left[p_k(700)\right] = (0.40)(0.05)$$
$$\left[p_s(300)\right]\left[p_k(600)\right] = (0.52)(0.04)$$
$$\left[p_s(400)\right]\left[p_k(500)\right] = (0.60)(0.03),$$ and so forth.

For this example, Table II shows that for the range of payload between 500 and 1500 lb, the value of P varies between 0.01 and 0.07, and that the optimum w_p varies from 200 to 400 lb. Table II also shows the best allocation of weight to guidance and warhead for each level of W.

To find each optimum division consider, for example, that the missile payload is to be 1200 lb. For this case one would enter the table at W = 1200 lb. and read from the P(W) column that the maximum P(1200) = 0.04, and this is obtained using w_p = 400. The remaining 800 lb is to be divided among w_w and w_g. Entering the table again with W = 800 lb. the optimum mix of w_w and w_g is read from their columns and is seen to be

$$w_w = 200 \text{ lb}$$

$$w_g = 600 \text{ lb}$$

This information is presented on Figure 3 for the entire range of missile payloads examined.

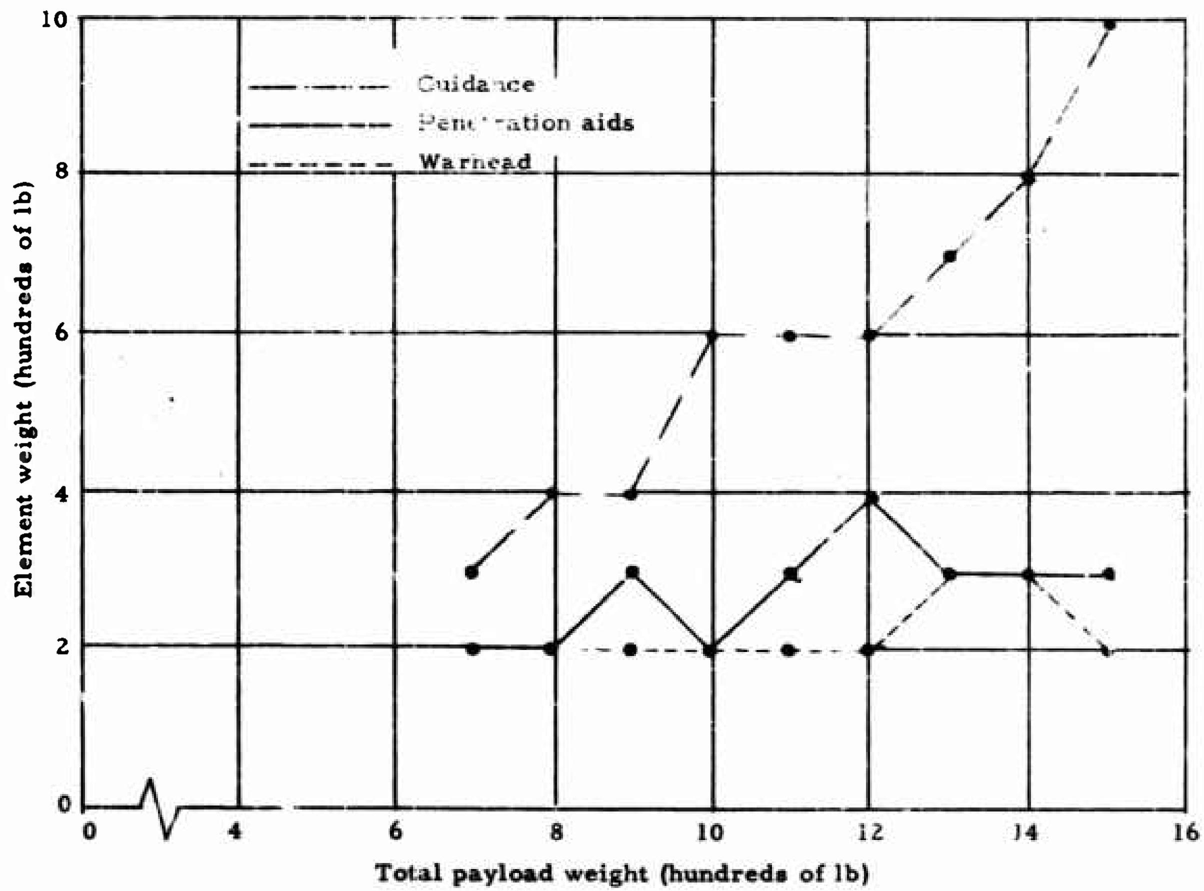


FIGURE 2. OPTIMUM DIVISION OF PAYLOAD FOR POINT TARGET DESTRUCTION
(Single warhead per missile and single missile per target)

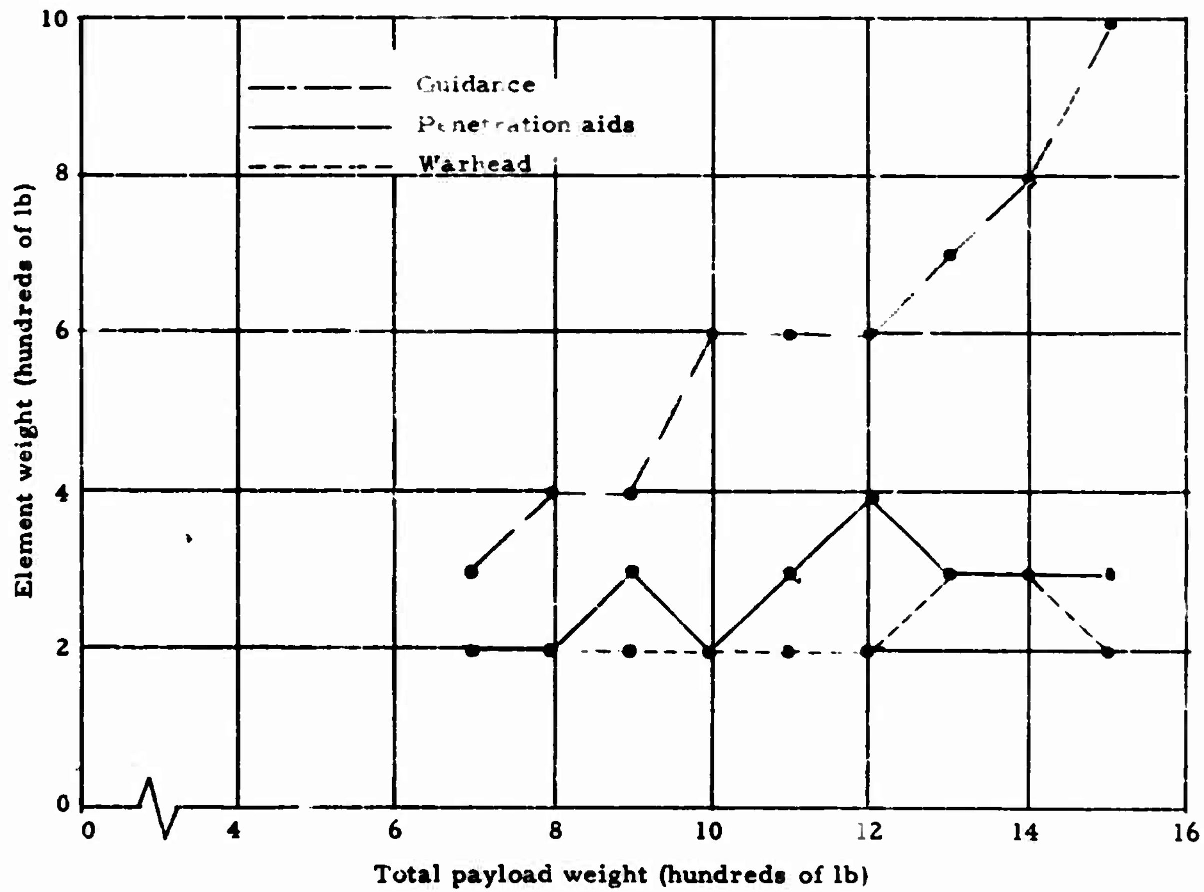


FIGURE 3. OPTIMUM DIVISION OF PAYLOAD FOR POINT TARGET DESTRUCTION
(Single warhead per missile and single missile per target)

5.0 LIMITATIONS ON UTILITY OF METHOD

The method presented was developed to solve a specific set of problems. It is essentially a simple method and given the data required, will afford quantitative results for the optimization criteria considered. But, as was pointed out, the analysis is based upon several design and operational considerations, the very nature of which will restrict the utility of the method for design purposes. First, because of design, development, and emplacement time and costs, it appears reasonable to expect that all missiles of a class will be equipped with identical warheads, guidance packages and penetration aids. On the other hand, it may be unreasonable to expect that all the targets for these missiles will have the same vulnerability and defenses. A design that is optimum for, say, the employment of a single missile against one target combination of hardness and defense capability may not be optimum for the employment of, say, two or three missiles against another target combination. A logical compromise might be, however, to choose the design that is optimum for anticipated employment against the most important set of targets and which also retains a high capability for other targets. The method of this example would be useful in this design compromise context.

Secondly, the guidance accuracy is, in general, dependent upon the range to target, and all targets for a class of missiles are certainly not at the same range. Here again, compromises would be necessary if this method is used.

A more detailed analysis could possible be employed to take account of these many intractable design and employment conditions. For example, an analytic method probably could be developed that would consider the use of the proposed missile against a large group of targets of varying worth, defense strength, and vulnerability. In the light of the problems raised above and during the analyses, however, it is not clear that a more detailed analysis is warranted. The type of design decisions considered here would need to be made early in the R & D program for a missile, and would therefore be based on early equipment (e.g., what will be the achievable CEP for a given weight and range to target) and intelligence (e.g., what defenses will the enemy employ for each target) estimates and early estimates of anticipated employment (e.g., how many missiles will be employed against each target and with what timing). Each of these could change substantially before the missile became operational, and the design that was optimal early in the R & D program would ultimately become only a compromise.

Perhaps, then, the greatest worth of a pre-design analysis using this method or any similar method, is that it would focus attention on the influence of the several required design, employment and intelligence estimates on the optimum payload division. A quantification of this influence and an analysis of the sensitivity of the design to the range of estimate uncertainty could serve as a useful point of departure for design compromises. Depending on the degree of estimate uncertainty, a sensitivity analysis could strengthen the apparent utility of any particular set of design parameters. Fortunately, the number of variables employed in this analysis is sufficiently small that the effects of uncertainty in a particular estimate could be clearly seen.

EXAMPLE D

OPTIMIZING A PRELAUNCH CHECKOUT[1/]

[1/] Condensed from "MISSILE PRELAUNCH CONFIDENCE CHECKOUT: CONTENT AND EQUIPMENT DESIGN CRITERIA," by S. I. Firstman and B. J. Voosen, RAND MEMORANDUM RM-2485-PR, February 22, 1960.

ABSTRACT

This example presents a procedure for determining the optimum test content of an ICBM prelaunch checkout that is subject to a time constraint. Cost considerations are not introduced as a constraint, but instead are employed after the text content has been optimized for each possible test duration constraint in order to select between designs. An example is given and references are cited that contain an explanation of the estimation of the parameters associated with the design technique.

TABLE OF CONTENTS

ILLUSTRATIONS

TABLES

1.0 INTRODUCTION

Most complex weapons are checked prior to their employment so as to "be sure" that the weapon contains no mission-failure-causing defects and is in fact ready to perform its assigned mission. Though often partially qualitative in nature, this check is intended to establish "confidence" in the ability of the weapon to perform its mission. Once a decision to employ a "confidence" checkout has been made, engineering and operational evaluations can be employed to determine the content of such checkouts and the required coverage of the checkout equipment. This example is concerned with the problem of making design decisions for time-limited checkouts of ballistic missiles and develops (a) quantitative criteria for the value of inclusion of individual checks in the checkout, (b) an expression for the required efficiency of the checkout equipment, and (c) a mathematical method for determining the content of a theoretically optimum checkout. A brief example is included. The method developed can also be used to ascertain the launch "confidence" of existing systems.

The performance of a weapon is a function of its design and its construction, factors that also result in an inherent flight reliability; one cannot "check" performance or reliability into a weapon. For many operational situations the only reasonable purpose of a prelaunch check is to assure the commander that the weapon is in a design or mission-ready condition at the time of employment. For this reason, the term "commander's confidence" is defined as: the probability that a missile, for which a launch attempt is made, does not contain a mission-failure-causing defect, or, equivalently, the probability that no undetected mission-failure-causing defects are present in the weapon after the prelaunch checkout.

This definition will be formulated mathematically, as a function of (a) the probability of a mission-failure-causing defect occurring in the weapon, (b) the capability of the checkout to detect each defect, and (c) the possible deleterious effects on the weapon of the checkout. This formulation, then, explicitly accounts for the tradeoff between detecting defects that are present prior to checkout, and of introducing defects during the checkout. The resultant probabilistic model employs estimates that are obtainable in missile reliability programs.

The object of a "confidence" check is to maximize the commander's confidence. In general, if all functions of a weapon could be checked prior to employment, this confidence would be a maximum. However, for most weapons, it is physically impossible and operationally impractical to check all functions of the weapon prior to employment. Ordinarily, for ballistic missiles the checkout will need be done within a limited time; sometimes with a limited volume of equipment, or within some other physical constraint. The time constraint is usually dominant, however, and therefore, the problem addressed is that of choosing which of the many tests to perform within the limited time allowed for such prelaunch checks, so a to maximize the confidence.

2.0 DESCRIPTION OF MODEL

The first step in developing the model is to describe the missile characteristics which affect the checkout decision. To do this, observe that a missile is composed of a group of physical functions (power supplies, engine controls, fuel pressurization equipment, etc.). At some level these functions can be grouped into a set of independent functions. It is desirable of course, to group these functions into logical test units. Many functions are dependent, i.e., are connected and operated such that an error in one propagates to other functions. Independence, of course, means that each function, or "black box" or group of black boxes, will fail but will not affect other functions when it fails.

As an example of independence, consider a power supply and a transmitter. They are certainly functionally dependent, but each can fail independent of the other and could be employed as independent functions in this model.

For operational missiles, the only defects that must be considered for the prelaunch checkout are those that: (a) could cause a mission failure and (b) can be feasibly checked within the operational environment. It is possible that some functions are not critical for mission success, and certainly, some functions, such as separation bolts, cannot be checked in an operational environment. Using standard reliability estimating techniques, the following set of probabilities can be estimated for each of the missile's defects which number, say, R, $r = 1, 2, \ldots, R$.

p_r = Probability that the rth defect does not exist prior to checkout, i.e., has not occurred since last checkout.

q_r = Probability that the rth defect is detected given that it existed prior to the checkout, and was checked.

r_r = Probability that the rth defect is not caused by being activated (or exercised) and not checked, given that it did not exist prior to checkout.

s_r = Probability that the rth defect is not caused by being activated and checked, given that it did not exist prior to checkout.

Examination of these probabilities will show that:[2/]

(a) p_r is related to the system reliability, and is simply the probability that the rth defect has not occurred since the last checkout of the missile. It is dependent on the elapsed time since the last checkout.

(b) q_r can be called the efficiency of the checkout equipment, and

(c) the primary difference between r_r and s_r is the level of operating stress to which the defect is subjected. For many circuits this distinction may be trivial. Both r_r and s_r are dependent on the time required for the complete checkout, while s_r is also dependent on the method of test employed and the time required for the test.

These probabilities can be combined into a "confidence" function that

---

2/ Reference 2 contains both a description of methods of estimating these terms and a quantitative treatment of the sensitivity of the decision process to errors in these estimates.

will estimate the probability that the weapon is in a design or mission-ready condition when a launch attempt is made, i.e., the probability that no undetected defects exist in a weapon after checkout.

If, during the checkout routine,

(a) a check is made for the rth defect, then

$C_r^{(1)} \equiv$ Probability that the rth defect (which is checked) either does not exist after checkout or has been detected

is by decomposition

$$C_r^{(1)} = (1 - p_r) q_r + s_r p_r + (1 - s_r) p_r q_r \tag{1}$$

$$= q_r (1 - p_r s_r) + p_r s_r \tag{2}$$

This is the sum of the mutually exclusive ways in which the rth defect, which is checked, will either not exist after checkout or be detected if it exists. The first term in the first equation is the probability that the defect existed prior to checkout and was detected; the second term is the probability that the defect did not exist prior to checkout and was not caused by being checked; and the third term accounts for defects being caused during checkout and being detected. It is assumed that the same q_r is applicable to both situations.

(b) a check is not made for the rth defect, but it is activated (turned on and/or exercised), then

$C_r^{(2)} \equiv$ Probability that the rth defect (which is not checked but activated) either does not exist after checkout or has been detected

is given by

$$C_r^{(2)} = (1 - p_r) a_r + r_r p_r + (1 - r_r) p_r a_r \tag{3}$$

$$= a_r (1 - p_r r_r) + p_r r_r \tag{4}$$

This is the sum of the mutually exclusive ways in which the rth defect, which is not checked but which is turned on during the process, will not exist undetected after checkout. The first term is the probability that a defect that exists prior to checkout is found indirectly. The second term is the probability that the defect is not caused during the checkout process, and the third term accounts for defects that are caused, but indirectly detected. It is assumed that the same a_r is applicable to both situations.

The quantity a_r is a dependent variable of the checkout process. It is defined as

a_r ≡ Probability that the rth defect is detected indirectly, given that it existed prior to the checkout,

and is assumed to be given by the dichotomy,

$$a_r = \begin{cases} 1, & \text{defect manifests itself in an overt manner that will be seen during the check of the functions that are checked.} \\ 0, & \text{defect does not manifest itself in an overt manner that will be seen during the check of the functions that are checked} \end{cases}$$

The value of a_r will depend on what functions of the system are checked, the checkout method, and system design.[3/]

Or,

(c) a check is not made for the rth defect and it is not activated,

then

$C_r^{(3)}$ ≡ Probability that the rth defect did not exist prior to checkout, is simply,

$$C_r^{(3)} = p_r \tag{5}$$

If x_r and y_r are defined as:

$$x_r = \begin{cases} 1, & \text{rth defect checked} \\ 0, & \text{rth defect not checked} \end{cases}$$

3/ To better understand a_r, consider a check of a radio transmitter. If, when the transmitter is turned on, the tubes don't light, or no power is emitted, one can deduce that either the transmitter or the power supply (including the linking cables) is at fault. A second check would be required to determine which function is faulty, but $a_{\text{power supply}} = 1$, in this case, if the transmitter is checked, because if the power supply is inoperative, a check of the transmitter will reveal this condition.

and

$$y_r = \begin{cases} 1, & \text{rth defect activated (not checked)} \\ 0, & \text{rth defect not activated (not checked)} \end{cases}$$

then the confidence function, C, the probability that none of the independent defects exist undetected after checkout, is the product of the $C_r^{(i)}$ above, that is pertinent to each defect r. By definition, assuming independence of terms,

$$C = \prod_{r=1}^{R} [q_r(1-p_r s_r) + p_r s_r]^{x_r} \cdot [a_r(1-p_r r_r) + p_r r_r]^{y_r} (p_r)^{(1-x_r-y_r)} \tag{6}$$

By taking the natural logarithm of C, the following expression is obtained:

$$\ln C = \sum_{r=1}^{R} [x_r \ln(q_r(1-p_r s_r) + p_r s_r) + y_r \ln(a_r(1-p_r r_r) + p_r r_r) + (1-x_r-y_r)\ln p_r] \tag{7}$$

As all functions in a missile must be activated (with the possible exception of some fuzing equipment in the re-entry vehicle and items such as stage separation mechanisms) prior to launch, a function may be checked or not checked, but it must be turned on prior to launch; therefore,

$$x_r = 1 \Rightarrow y_r = 0$$

$$y_r = 1 \Rightarrow x_r = 0$$

and moreover

$$x_r + y_r = 1$$

Employing the third relationship between x_r and y_r (i.e., by observing that $y_r = 1 - x_r$), the natural logarithm of C becomes 4/

$$\ln C = \sum_{r=1}^{R} \left(x_r \ln \left[\frac{q_r (1 - p_r s_r) + p_r s_r}{a_r (1 - p_r r_r) + p_r r_r} \right] + \ln (p_r r_r + a_r (1-p_r r_r)) \right) \tag{8}$$

4/ An implicit assumption on the nature of r_r is required for this formulation. It is assumed that r_r is the same for all situations involving check of associated systems.

As stated, for a ballistic missile prelaunch checkout, the intent is to maximize C, and hence ln C, subject to the single dominant constraint, which is time. This means that the "confidence" maximization problem is to choose a set of $x_r = 1$, such that ln C is a maximum and $\sum_r x_r t_r \leq T$, where T is the time allowed for checkout, and t_r is the time required to perform the rth check.

3.0 TECHNIQUE OF SOLUTION

The first step in the process of deciding which x_r become 1 and which remain at zero, is to observe the nature of the ratio,

$$\frac{q_r(1 - p_r s_r) + p_r s_r}{a_r(1 - p_r r_r) + p_r r_r}$$

This is the ratio of

$$\frac{\text{P [rth function contains no undetected defect if checked]}}{\text{P [rth function contains no undetected defect if not checked]}}$$

If this ratio is ≤ 1, then its contribution to ln C is ≤ 0. Moreover, if the ratio is < 1, then a check of the rth function (or defect) could do more harm to the system than good. This situation could arise in missiles, where the prelaunch life of the missile is spent in essentially a "shelf" or non-operative condition, and p_r could, for many systems (especially mechanical systems such as rocket engines) be essentially unity.

By this reasoning, then, it can be determined that the rth function (or defect) should not be checked unless

$$q_r (1 - p_r s_r) + p_r s_r > a_r (1 - p_r r_r) + p_r r_r \tag{9}$$

For $a_r = 0$, the above criterion means that before one would consider checking the rth function, the checkout efficiency for the rth function, q_r, should be

$$q_r > \frac{p_r}{1 - p_r s_r} (r_r - s_r) \tag{10}$$

For $a_r = 1$, in a similar manner, q_r must be such that

$$q_i > \frac{a_r (1 - p_r r_r) + p_r r_r - p_r s_r}{1 - p_r s_r} \tag{11}$$

which gives

$$q_r > 1 \tag{12}$$

which is physically impossible.

These two bounds on q_r can both be employed. For $a_r = 0$, which means that in order to find the rth defect, an explicit check must be performed, q_r given by Equation (10), is a lower bound on the checkout efficiency and can be employed as a design criterion for the checkout equipment. In order for the checkout equipment to be fruitfully employed, it must be at least this good.

For $a_r = 1$, which means that because of the other checks that are being performed, the rth defect will be detected (if it exists), the lower bound on q_r was seen to be $q_r > 1$. This is clearly impossible and the reason for this result is apparent. If the existence of the defect can be detected implicitly, there is no reason to make an explicit check for it. This means that

$$a_r = 1 \Rightarrow x_r = 0$$

is a valid condition.

Employing the above condition on a_r and x_r makes it possible to view the problem of what to check so as to maximize ln C as a two-stage time allocation problem. The first stage will determine which tests are best performed if each must be done explicitly, then the second stage will take account of the dichotomy between a_r and x_r to improve the initial solution.

4.0 FIRST STAGE OF ALLOCATION

First, for only those r for which the checkout equipment can achieve a checkout efficiency as given by Equation (10), it shall be assumed that all $a_r = 0$ and

$$Z_1(T) = \ln C - \sum_r \ln p_r = \sum_r x_r \ln\left[\frac{q_r(1-p_r s_r) + p_r s_r}{p_r r_r}\right] \quad (13)$$

shall be maximized subject to

$$T \geq \sum_r x_r t_r \quad (14)$$

Observe the nature of the term in brackets. For a particular function/defect, and for a particular operational philosophy and test concept, this term is merely a parameter. The problem as stated, then, is to maximize a linear function, subject to a single linear constraint.

The problem of choosing which $x_r = 1$ is now of linear programming (LP) form, and as just one constraint is being employed, the problem becomes an LP problem of the Knapsack type[1] and is readily solved in graphical form. For the defects $r = 1, 2, \ldots, R$, each of which has an associated $\ln[\;]_r$, and t_r, plot the set of t_r, $\ln[\;]_r$ as follows:[5/]

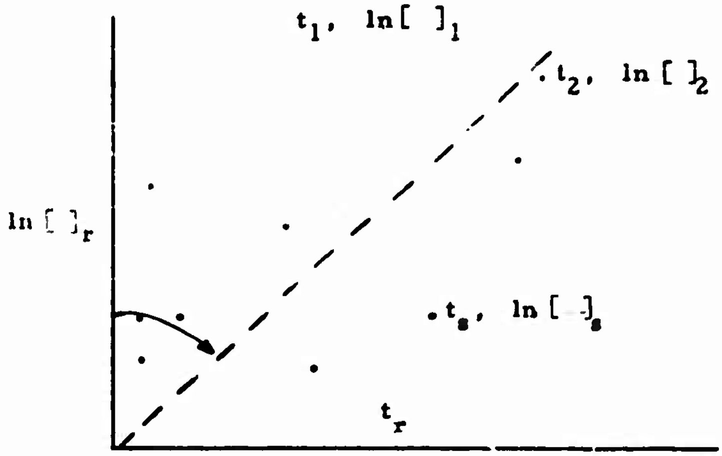


FIGURE 1. PLOT OF $\ln[\;]_r$ VS t_r

5/ The notation $\ln[\;]_r$ will be employed instead of the cumbersome term $\ln\left[\frac{q_r(1 - p_r s_r) + p_r s_r}{p_r r_r}\right]$

Now, the problem is to select a subset of these points that represent the set of defects (or functions) that are to be checked in a time-limited checkout. Then, as done by Dantzig,[1] rotate clockwise a ray with the origin as pivot point and $\ln [\ ]_r$ axis as starting position. Tests corresponding to points swept out by the ray are selected in turn until the sum of their test times meet or just exceed the time limitation. If for the jth test the time limitation would first be exceeded, a decision must be made as to whether that test should be conducted and the checkout time lengthened by a small amount, or, whether that test should be excluded.

The logic of this decision process is simply that those tests with the greatest value per unit time are the first chosen for the set, i.e., those tests with the largest value of $\ln [\ ]_r / t_r$ are chosen until $\Sigma\ t_r = T$. To these tests chosen by r_c, then, the value of $Z_1(T)$ becomes

$$Z_1\ (T) = \sum_{r_c} \ln \left[\frac{q_r(1 - p_r s_r) + p_r s_r}{p_r r_r} \right] \tag{15}$$

5.0 SECOND STAGE OF ALLOCATION

The second stage in the time allocation process is to determine how the confidence can be improved by letting some $a_r = 1$, i.e., by getting some test information free. The second stage, then, is concerned with maximizing the entire confidence function, and can be represented symbolically by the quasi-functional equation

$$Z_2\ (T) = \underset{0 < t_2 < T}{\text{Max}} \left[R_2\,(t_2) + Z_1\ (T - t_2) \right] \tag{16}$$

where

$$R_2\,(t_2) = \sum_r \ln \left[a_r(1 - p_r r_r) + p_r r_r \right] \tag{17}$$

where t_2 is time allocated to second stage.

For each level of T, implicit time will be assigned to the second stage, $R_2(t_2)$, so as to maximize the sum of the two stages. This means that if, say, an $a_r = 1$ is brought in, the corresponding $x_r = 0$, and another check could possibly then be performed in the time which was previously allocated to the rth check.

It will be advantageous to allow an a_s to become one, if during the first stage the corresponding x_s became one, only if the other check, say the tth check, which could then be performed in the time previously allocated to the sth check, would cause the combined contribution of the sth and tth check to be greater after a_s and x_t became 1. For the cases when $t_s \geq t_t$, see Figure 2.

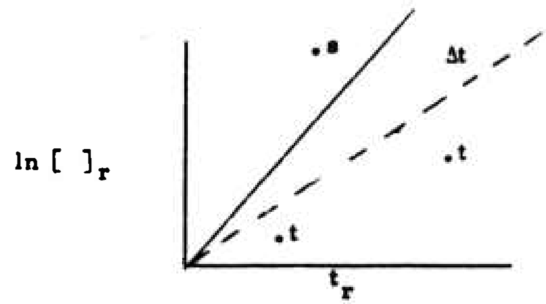


FIGURE 2. PLOT OF ln $[\,]_r$ VS t_r WHEN $t_s \geq t_t$

It can be shown that[6/] it is always advantageous to let $x_t = 1$ and obtain the benefit of $a_s = 1$.

Similarly, for the third case, when $t_t > t_s$, let $x_t = 1$ and $a_s = 1$ if, and only if,

$$f(\Delta t) < \ln \frac{q_t (1 - p_t s_t) + p_t s_t}{p_t r_t} - \ln (q_s (1 - p_s s_s) + p_s s_s), \qquad (18)$$

where $f(\Delta t)$ can be viewed either as the loss due to moving the ray counter-clockwise or the gain which could otherwise be obtained by the use of $t_t - t_s$. This function is readily tabulated for each physically possible such combination of s and t.

The preceding statements of conditions for substituting test t for test s in the set of $x_r = 1$, will also hold if several tests are required to allow some $a_s = 1$.

6/ See Reference 2.

The results of the preceding analysis allow the second stage of the allocation problem to be solved in an iterative manner. Given the plot of ln [$]_r$ vs. t_r, the iterative procedure is:

1. Look for any conditions of dependence that have already been satisfied. If some $x_t = 1$ and this allows an $a_s = 1$, where both s and t are contained in the set of $x_r = 1$, then $a_s = 1$, and move the ray down by an amount t_s.
2. Repeat step 1 until no more such situations can be found.
3. Search for other defects which have property $x_t = 1 \Rightarrow a_s = 1$, such that $t_t \leq t_s$, and s is of the set $x_r = 1$ and t is of the set $x_r = 0$. Substitute the tth test for the sth test and move ray down by an amount $t_s - t_t$.
4. Repeat steps 1, 2 and 3 until no more such situations can be found.
5. Search for other defects which have property $x_t = 1 \Rightarrow a_s = 1$, such that $t_t > t_s$, and s is of the set $x_r = 1$ and t is of the set $x_r = 0$. If, as previously defined, $f(\Delta t)$ exists, or potentially exists, and if Equation (18) holds, then let $x_t = 1$; otherwise, leave $x_t = 0$.
6. Repeat steps 1, 2, 3, 4 and 5 until no more such situations can be found.

6.0 CONDITIONS OF OPTIMUM SOLUTION

This iterative procedure will lead to a solution that can readily be seen to be optimum under three situations; either (a) all conditions of dependence have been utilized, such that all possible $a_r = 1$, or (b) those that have not been utilized are obviously not permissible, or (c) the use of those not utilized would obviously not improve the confidence.

Under conditions other than these three extremes, the resultant set of $x_r = 1$ cannot be proven optimum, and indeed may not be optimum due to the possibility that some unique combination of tests could be better employed. In this case, the answer obtained is at least a good and useable approximation to an optimum solution.

Each p_r will decrease independently as a function of time since last checkout, while q_r, r_r, s_r and will, in general, remain constant for a given test method and equipment design. Therefore, especially for electronic systems, it would probably prove profitable to consider several levels of time-since-last-checkout when using the model to see if any changes in optimum routine will occur. These would require a design compromise, and the method of handling such possibilities is discussed in Reference 2.

7.0 APPLICATION

This section demonstrates the application of the checkout model to a hypothetical single-stage missile that is based on designs of several present ballistic missiles. Although the illustration has been confined to less than 150 functions, all major functions of an actual operational missile have been included; i.e., a radio guidance system, control system, power supply, etc.

7.1 General Procedure

There are four general steps for determining what to check.

1. Aggregate missile functions into independent functions/defects using a functional block diagram of the system.
2. Obtain the necessary probabilities p_r, r_r, and s_r from reliability information, and from the ground support equipment designers obtain the time to perform each check, t_r, and the estimated efficiency, q_r. [7/]
3. Evaluate $\ln\left[\dfrac{q_r(1 - p_r s_r) + p_r s_r}{p_r s_r}\right]$ for each test unit.
4. Plot $\ln[\ ]_r$ versus t_r and determine the functions to be checked within the time constraint, as described in the

---

7/ Reference 2 contains discussion of methods to estimate each of these terms.

preceding section. [8]

This procedure is applicable to all functions that are candidates for use in the model. Some functions will be handled outside the model. For instance, there may be a military requirement to check a destruct system or some "fail safe" system, regardless of its reliability. These items obviously would be handled outside the model. The time to check these functions would be subtracted from the total time allotted for the confidence check. The remainder would be the time available for checks of the defects.

7.2 Independent Defects

As previously stated, the functions/defects considered must be independent. This requirement can usually be satisfied at some level of aggregation. If possible, however, aggregations should be made consistent with the selection of test units. A test unit will usually be what is commonly referred to as a "black box" or series of black boxes.

In one respect, all functions of a missile are dependent since the failure of any part of the basic system will cause the mission to fail, or at least will degrade the effectiveness. Most functions are independent, however, in a failure sense. For instance, a failure in an amplifier channel will not be dependent upon a failure in the gyro. If, however, the failure of an amplifier channel would cause a failure of some other function, say the hydraulic servo, the failures would be dependent and should be aggregated into a single defect.

As an example, Figure 3 is a copy of part of an actual block diagram of the control system for a current ballistic missile. The servo amplifier, hydraulic servo, and follow-up potentiometer are considered dependent in a failure sense because the design is such that a failure in any one function will cause a failure in the other functions. Therefore, the functions are aggregated in Figure 4 as Defect 117 (see Figure 4 for defect coding system for the entire missile). In contrast, a failure of the yaw rate gyro (107)

---

[8] The sum of the $[\ \ ]_r$ terms is plotted for tests that can be done simultaneously.

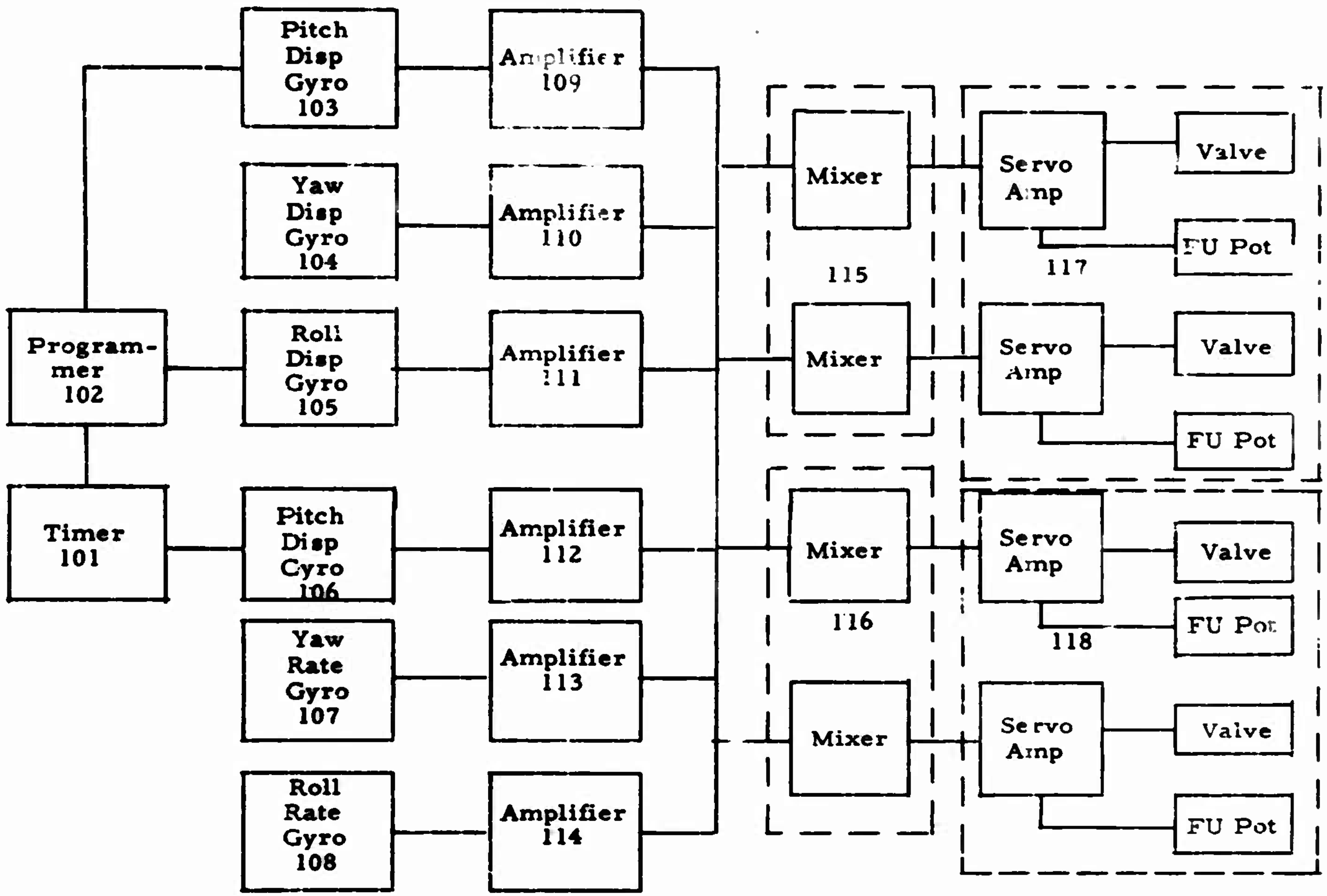


FIGURE 3. FLIGHT CON[illegible] SYSTEM BLOCK DIAGRAM (PARTIAL)

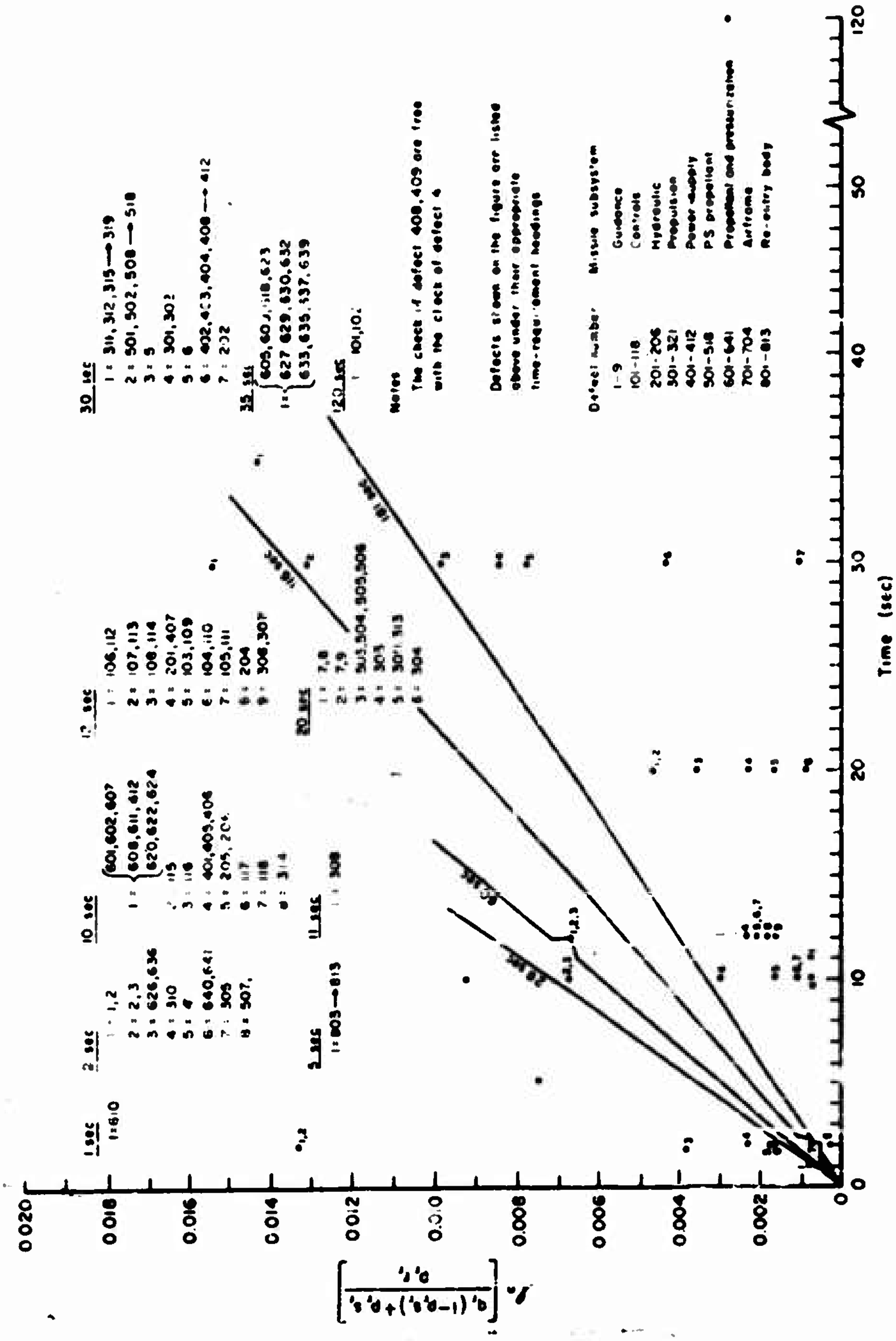


FIGURE 4. DECISION GRAPH
(1 WEEK AFTER PERIODIC)

will not cause a defect in the amplifier channel (113) if checked separately. Therefore, the gyro failure is considered to be independent of other failures.

One method for determining the dependence or independence of each function is to consider separately for each function on the functional block diagram, whether if it fails, will that failure cause a failure in some other function? If the answer is yes, then the functions must be aggregated in this model.

Using this type of analysis, the functional block diagram of each of the systems under consideration can be redrawn into independent defects. A comparison of the solid and dashed lines on Figure 3 shows that for this control system there was little change. All systems used in the illustration required about the same amount of aggregation. However, they will not be shown in this brief example.

7.3 Determining What To Check

After the data has been accumulated, the parameter $\ln [\ ]_r$ can be computed for each function and plotted against the parameter time for check, t_r. Table I contains the data for the flight control system for p_r, based on four weeks since the last periodic. This data and similar data for the balance of the missile's subsystems are plotted on Figure 4. As previously described, by rotating a ray clockwise, checks are selected in turn until the sum of the checkout time meets the time limitation. Since only those tests with the greatest "value" per unit time are chosen, confidence will be maximized. The one, two and three minute "best" checkout times are shown on Figure 4.

7.4 Design Decision Context

It is anticipated that this model and decision process will prove useful as a check on existing designs, to ascertain the estimated confidence afforded by existing designs, and also as a preliminary design tool. This design utility is envisioned as occurring in a process as follows (see Figure 5).

A decision is made to perform a prelaunch checkout on a particular

Defect No.	Description	p_r	q_r	r_r	s_r	a_r	t_r	$\ln [\ ]_r$
101	Timer	0.9993	0.95	0.9990	0.9989	0	120 sec	0.001611
102	Programmer	0.9993	0.90	0.9995	0.9995	0	w/101	0.001080
103	Pitch Disp Gyro	0.9993	0.80	0.9994	0.9994	0	12 sec	0.001040
104	Yaw Disp Gyro	0.9993	0.80	0.9994	0.9994	0	12 sec	0.001040
105	Roll Disp Gyro	0.9993	0.80	0.9994	0.9994	0	12 sec	0.001040
106	Pitch Rate Gyro	0.9946	0.80	0.9982	0.9980	0	12 sec	0.005737
107	Yaw Rate Gyro	0.9946	0.80	0.9982	0.9980	0	12 sec	0.005737
108	Roll Rate Gyro	0.9946	0.80	0.9982	0.9980	0	12 sec	0.005737
109	Amp Channel	0.9993	0.90	0.9996	0.9996	0	w/103	0.000990
110	Amp Channel	0.9993	0.90	0.9996	0.9996	0	w/104	0.000990
111	Amp Channel	0.9993	0.90	0.9996	0.9996	0	w/105	0.000990
112	Amp Channel	0.9993	0.90	0.9996	0.9996	0	w/106	0.000990
113	Amp Channel	0.9993	0.90	0.9996	0.9996	0	w/107	0.000990
114	Amp Channel	0.9993	0.90	0.9996	0.9996	0	w/108	0.000990
115	Mixer	0.9946	0.90	0.9979	0.9976	0	10 sec	0.006738
116	Mixer	0.9945	0.90	0.9979	0.9976	0	10 sec	0.006738
117	Servo Amp	0.9993	0.90	0.9995	0.9995	0	10 sec	0.001080
118	Servo Amp	0.9993	0.90	0.9995	0.9995	0	10 sec	0.001080

TABLE I

FLIGHT CONTROL SYSTEM DATA

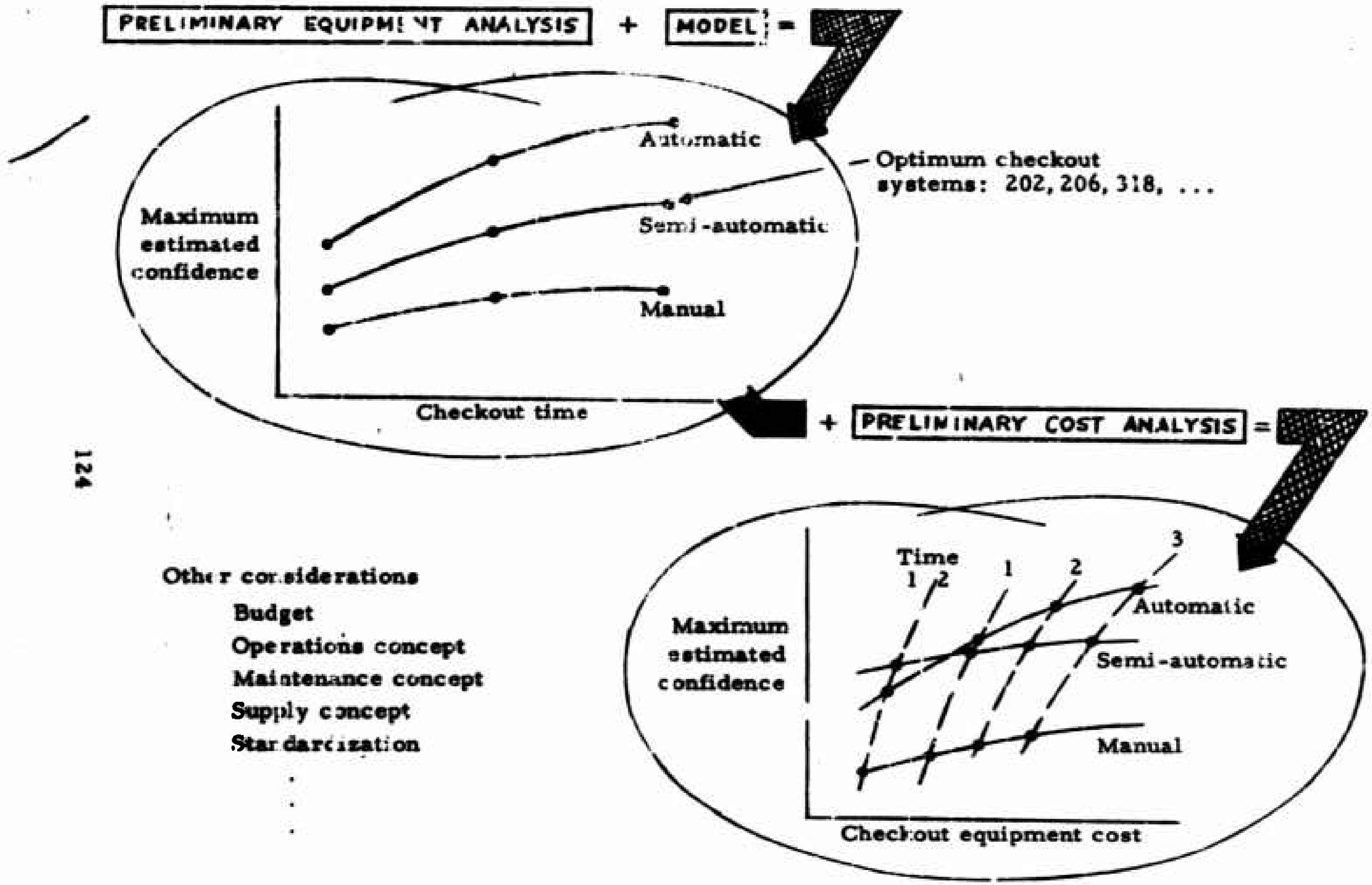


FIGURE 5. DESIGN DECISION CONTEXT

missile. Unknown, at this point in time, is the method by which the checkout will be performed, what missile capability the checkout will assure, and what checks will be done to obtain this capability.

The AGE designers are told to perform a comparative analysis of alternative means of performing the prelaunch checkout. For this hypothetical design process, three control means that are representative of three control and resultant design concepts are selected for study; these are digital computer, punched paper tape and manual. Each is representative of a control mechanism in a particular test logic and control realm.

A preliminary design analysis is conducted using each of the three design concepts. This preliminary analysis should consider: (a) the problems of how best to check each function within each design concept; (b) test point availability and added weapon weight due to additional test leads; (c) the equipment requirements (stimulus and measurement) for each function test; and (d) estimates of the relevant parameters for each test.

For several levels of checkout time (constraint) the checkout model would then be employed, using each design concept in turn, to choose the optimum set of tests to perform from the set of tests that are competitors for performance. Then using the confidence function the maximum confidence could be estimated for each combination of design concept and time.9/

This maximum estimate "confidence" can be plotted as in the upper

---

9/ It will probably prove true that this design process will be an iterative procedure. An analogous process is the weight allocation among systems of a missile. Initially, preliminary weight allotments are made, and designs are initiated with these weight restrictions in mind. After a first cut at the systems designs, a revised weight budget is considered. In a similar manner, the times to perform individual tests are initially estimated. If, after some design effort, it appears that some of the initial estimates were in error, the checkout content problem should be reconsidered using the revised time estimates. The same type of arguments can be raised for estimates of p_r, q_r, r_r, and s_r. This design modification procedure, it is realized, is feasible only until the design must be fixed. Thereafter, system modification becomes much more expensive, both in time and dollars, and these considerations could override any desired design changes.

graph of Figure 5. As indicated, the optimum checkout is known for each point.

Following this model exercise, a preliminary cost analysis would be performed. This would result in a plot of the data such as the lower graph of Figure 5.

Now a decision of what checkout system to design or procure can be made within the context of estimated confidence, time and dollars. The trade-offs between these elements and also between degrees of equipment automaticity are explicitly shown. Other considerations which can also be brought to bear upon the design and procurement problem are the operational concept, maximum budget, site manning philosophy, desire for standardization, etc. In this overall cost-logistics-operational context the design decision can be objectively made, and the checkout model is seen to be an integral part of this design and decision process.

It is seen that in this analysis, system cost is not used as a design constraint, but is employed after the design is optimized for each level of checkout time to choose which design to procure. In this manner the design is optimum with respect to the true operational constraint, namely time for checkout, and the effects of increasing or decreasing the budget for a given checkout time can be observed.

REFERENCES

(EXAMPLE D)

1. Dantzig, G. B., "Discrete-Variable Extremum Problems," P-876, The RAND Corp., December 12, 1956.

2. Firstman, S. I. and Vooren, B. J., "Missile Prelaunch Confidence Checkout: Content and Equipment Design Criteria," RM-2485, The RAND Corp., February 22, 1960.

EXAMPLE E

MISSILE AVAILABILITY

ABSTRACT

The availability of a system subjected to a sequence of calendar spaced checkouts is considered. Formulae for calculating the optimum frequency of checkout are given for the situation which considers checkout time as down time. Imperfect repair, imperfect checkout, and resource limitations are treated. A technique for the estimation of two parameters of the availability model is also given.

TABLE OF CONTENTS

ILLUSTRATIONS

TABLES

1.0 INTRODUCTION [1/]

In the not-too-distant past, when a new weapon system (usually a manned aircraft) was phased into the Air Force inventory, scheduled maintenance requirements were typically approximated by extrapolating from experience on previous similar systems. With few exceptions the indicator used was flying-hours, which usually gave a fair measure of the amount of activity a system had experienced; the necessary maintenance attention was predicated on this indication. More importantly, the flying program exercised all or nearly all systems and subsystems in a manner closely related to expected wartime actions, thus verifying the condition of the weapon system generally, and bringing defects to light when and where they existed.

The introduction of large, complex missiles, major portions of which are normally inert, has changed this picture considerably. Since these systems do not fly periodically, critical defects may remain hidden for unacceptably long periods of time unless appropriate means for identifying them are developed and applied. This section is concerned with one aspect of the problem, namely, the quantitative relationship between checkout frequency and readiness.

2.0 A SAMPLE PROBLEM

The problem of how often to "exercise" (i.e., verify the launch readiness of) a ballistic missile or its systems is a knotty one. On the one hand, an "exercise" may take the system off alert for the time required to perform it. The "exercise" may also cause a failure in a good system, and result in further downtime required to repair the system. These problems respond to a very simple solution--namely, fewer operational exercises.

[1/] The relationships of availability, dependability, and capability are discussed in Volume II of the TG-II Report.

On the other hand, undiscovered failures also have to be considered, because failures can occur while the system is standing in a ready condition, on alert. Many of these undiscovered failures can be found only by conducting operational exercises. Again the answer to the problem is simple: to cope with undiscovered failures, conduct more operational exercises; they will become evident and be corrected (though one or more new ones may occur in a short time). Clearly, these two simple solutions are mutually incompatible. In fact, the problem resolves into one of deciding where the optimum position is between too many and too few exercises.

We will restrict our attention here to systems whose failures occur exponentially. This characteristic failure pattern has been the subject of many discourses on statistical methods and models. (One of the more easily understood mathematical derivations for it appears in a RAND Research Memorandum by R. R. Carhart,(1) who notes that the words "random", "accidental", "chance", and "Poisson" have all been used to describe this type of failure distribution, whose formal designation is a negative exponential distribution of failures.)

Historically, it has been found that many components, subsystems, and systems experience exponential failures. The most detailed analysis was conducted by D. J. Davis,(2) who found that the exponential failure distribution characterized a wide variety of devices, from ball and roller bearings, vacuum tubes, and many other electronic systems and components, to passenger-bus motors and airborne radar systems. Boodman(3) presented further studies on airborne radars, and Lusser(4) showed information which added naval torpedoes to the list. More recently, the Atlas ICBM was shown to display this behavior.(5)

The operational conditions under which the negative exponential failure distribution occurs have likewise been a topic of considerable interest, and of some disagreement. Carhart noted that field data show exponential failures to be prevalent in equipment with (1) long periods of operation, (2) relatively constant operating conditions, and (3) continual replacement of failing components. He noted that vague explanations are usually offered for the first two of these characteristics. Boodman and Davis take issue with Carhart's second conclusion, suggesting that severe, unpredictable stresses characterize the exponential failure pattern. In any case, the exponential distribution of failures seems to represent quite well the behavior of those systems which have no gross wearout phenomenon in their operating lifetimes; and it seems to represent equally well those systems with discrete component-wearout phenomena, provided the components of the system are replaced as they fail, and that their ages thus become mixed after a time, in a representative system. (See Reference 6 for a proof.) The most striking example of this is probably the bus-motor case cited by Davis, in which (with all parts new at the outset) the elapsed time before the first major overhaul had a Gaussian distribution (i.e., with a distinct wearout pattern), while the third and subsequent overhauls (with part ages now mixed) occurred exponentially.

The exponential failure phenomenon has a crucial implication for maintenance policy; it says that preventive maintenance, in the normally accepted sense, is pointless.[7] A system, subsystem, or component which has an exponential failure characteristic, and which is known to be in operating condition, is just as likely to survive for a given time period as is a replacement which may be new or newly overhauled. Thus, periodic replacement of systems, modules, or components is not warranted unless it is known they will wear out within the anticipated system lifetime. Periodic inspection becomes equivalent to periodic renewal.

A second interesting characteristic of exponential failures is the attractive analytic behavior. All the important quantities can be represented analytically by linear equations or exponentials; and exponentials can often be approximated by linear expressions. This makes the exponential failure pattern an intriguing one for "paper" studies of many kinds, the most obvious of which is optimization in one form or another.

Our purpose here is to relate checkout frequency to the operational-readiness characteristics of static alert systems; that is, those systems which are kept in readiness for emergency use at some future time in a role of retaliation, defense, search, or any number of other functions. We will deal primarily with systems which spend most of their operational lives in an "off" or "standing" condition, although only minor modifications of the technique are necessary to treat those systems which operate continuously (or nearly so). We will deal with systems having an exponential failure distribution. In order to demonstrate clearly the method used and the technique for including various idiosyncrasies of particular operational situations, we will start with a very simple model, and then add complicating factors to it.

3.0 THE SIMPLEST MODEL

Consider a system, weapon, or part which is to be maintained in a standing condition (static alert) for a large number of time periods.[2/] The system has an exponential failure distribution with a mean time to failure t_f and a corresponding constant failure rate $\lambda = \frac{1}{t_f}$. The system receives a checkout every T time periods, which imposes a stress (probability of failure) q on a good system. The checkout time is negligible, and it is assumed that the checkout will uncover any existing failure, whether caused by the checkout or merely by standing. It takes an average of R time periods to correct a malfunction and the time-deterioration of the system starts anew at the end of the checkout.

Figure 1 shows an "out of commission" profile (the heavy solid line) for the situation described here. The ordinate is the probability of the system being out of commission as a function of time. The system, weapon, or part is assumed to be in commission at $t = 0$, and the probability of failure increases according to the exponential relationship. A failure during this period is initially undiscovered; that is, the operator is not yet aware of it. A checkout is performed at time T. At this time any failure becomes known, whether a previously undiscovered one or one caused by the

---

2/ For illustrative purposes, one day will be the time period used here, although it can be any other convenient unit, such as a minute, hour, week, or month. The need for a large number of time periods is discussed later in this section.

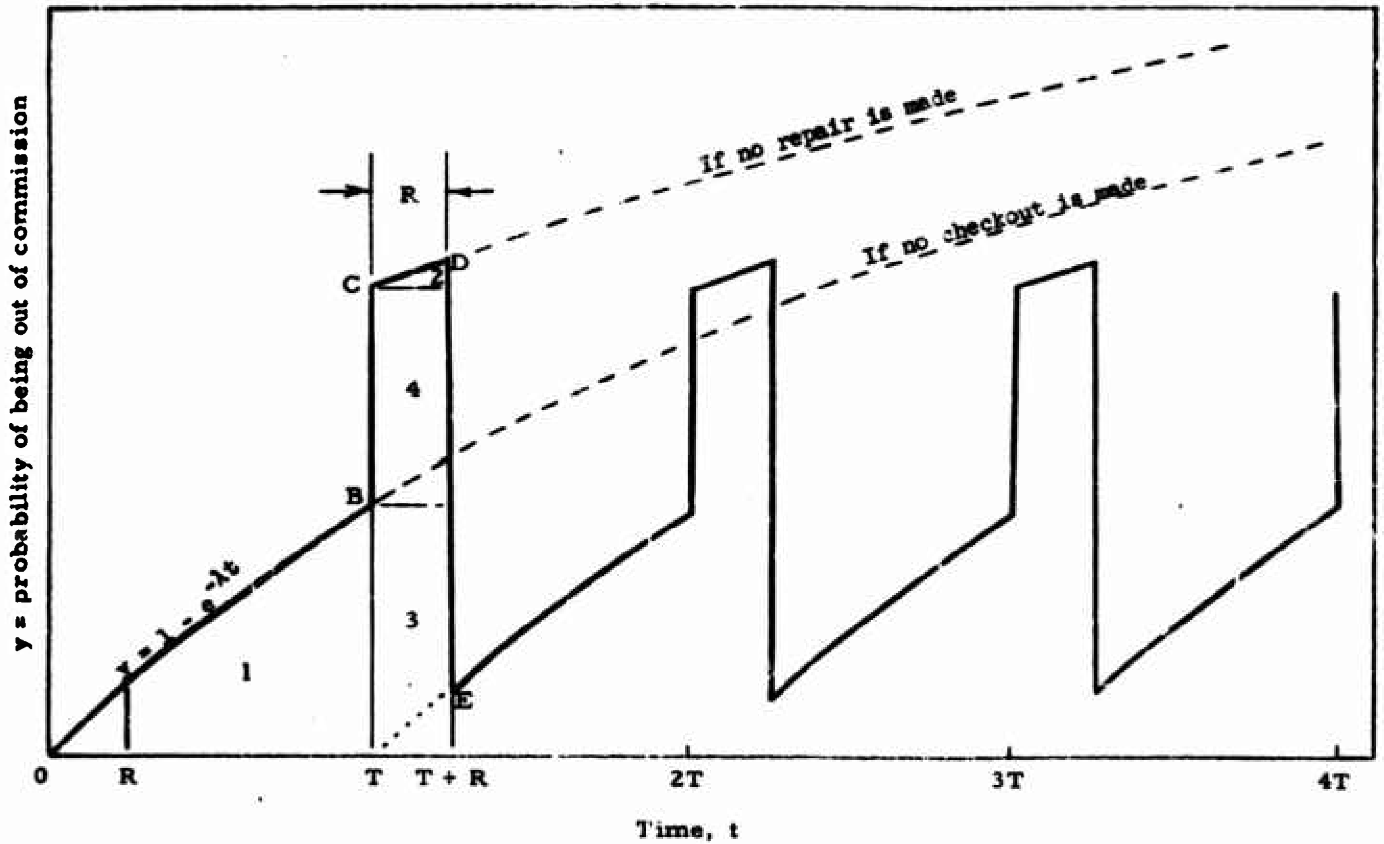


FIGURE 1

OUT-OF-COMMISSION PROFILE

checkout itself. The latter quantity is the product of the checkout failure probability q and the probability of the system's being "good" at time T (checkout cannot cause a failure which has already occurred). Thus, the line BC is of length $q [1 - (1 - e^{-\lambda T})]$, or simply $qe^{-\lambda T}$ If a failure is found, a repair is made or a replacement system is installed (taking a time R). The repaired or replaced system is not checked subsequent to installation. Thus, the original failure rate (with the system starting good) resumes at point C. However, line segment CD does not have the same slope as segment TE (shown dotted) since only a system which has passed the checkout (i.e., $[1 - q] e^{-\lambda T}$) must be considered as going out of commission (unknown to the operator) at the rate λ. A system which failed the checkout is in repair and should not be counted out of commission twice.

At time T + R, the repair or replacement process is complete and a system is in commission except for the probability (unknown to the operator) of having failed since the previous checkout, whether it was standing or in repair. From this point the process is repeated and is seen to be cyclic[3/] with period T after the initial transient caused by our assumption of a good system at t = 0.

Note that the ordinate y in Figure 1 accounts for all ways of being out of order, whether or not the operator is aware of the condition. Specifically, it accounts for a finite probability of failure in a system which has not been checked out for periods of time ranging from 0 to T, and which is usually considered in commission according to normal Air Force terminology. Thus, it is desirable to define here what might be called a "real" in commission rate or, preferably, "ready rate," to distinguish it from the other connotation. This will be the probability of a system's being in a good condition, when one accounts for the probability of an existing but undiscovered failure.

3/ Note that the cycle length is constant, a desirable arrangement if a checkout team is to go from system to system on a scheduled basis. For a similar derivation, when cycle length depends on the outcome of the checkout, see Reference 8.

If we wish to find an analytic expression for the average probability of a system's being in a good condition, we can find the area under the out-of-commission curve for one cycle, divide by the cycle length, and subtract from one. (This average will fairly well represent the ready rate for a group of such systems if checkouts of the systems are staggered instead of simultaneous.) The total area under the curve is most easily found by summing areas 1, 2, 3, and 4 in Figure 1:

$$A = \text{Area 1} + \text{Area 2} + \text{Area 3} + \text{Area 4} \tag{1}$$

Substituting the area equivalents, we have

$$A = \int_R^T (1 - e^{-\lambda t})\, dt + (1 - q)\, e^{-\lambda T} \int_O^R (1 - e^{-\lambda t})\, dt + (1 - e^{-\lambda T})\, R + qRe^{-\lambda T} \tag{2}$$

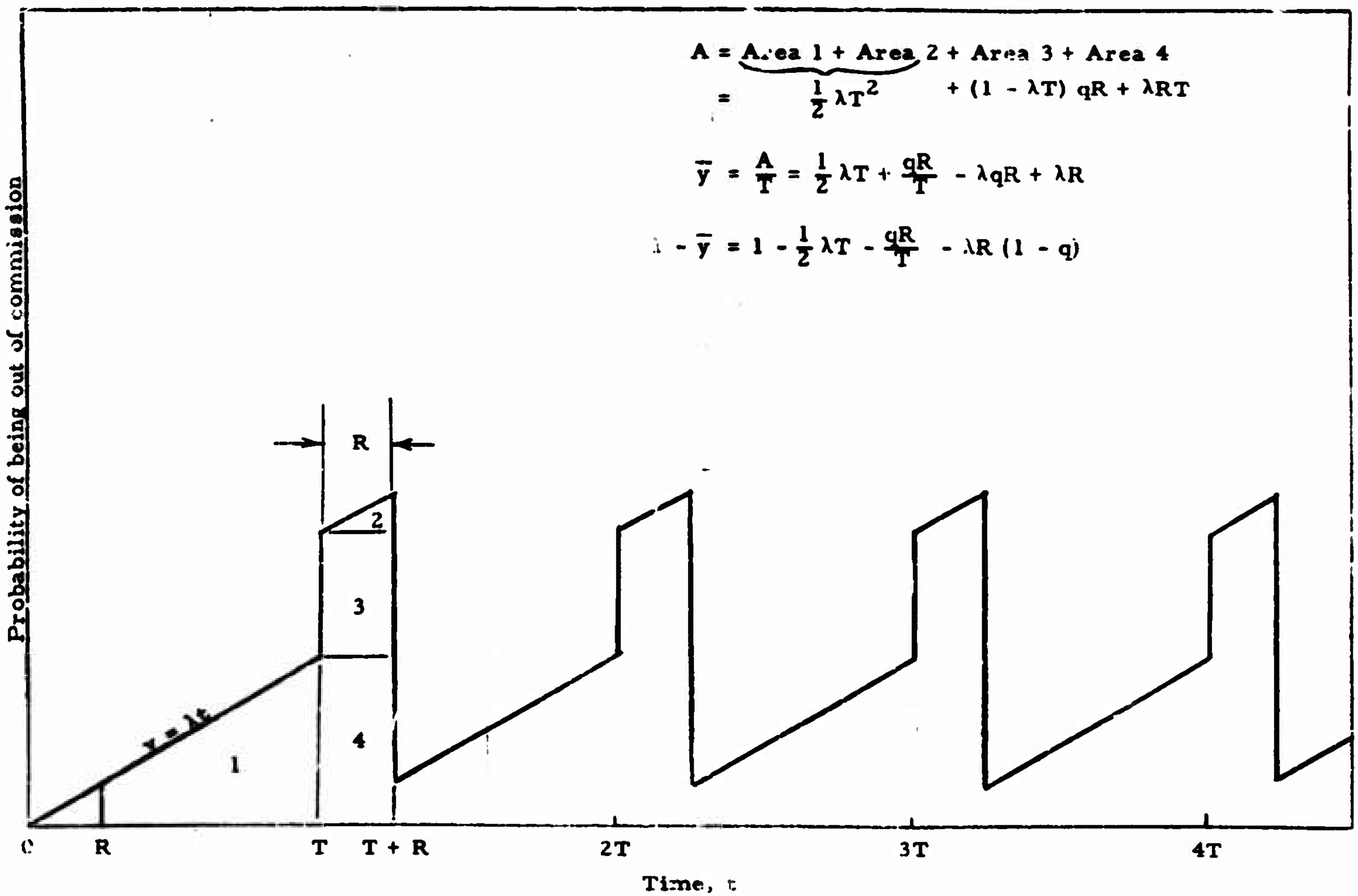
Integrating, cancelling like terms of opposite sign, dividing by the cycle length T, and subtracting from one, we find the proportion of time in commission (the ready rate) is:

$$1 - \bar{y} = 1 - \frac{A}{T} = \frac{e^{-\lambda R} - (1 - q)\, e^{-\lambda(T + R)} - qe^{-\lambda T}}{\lambda T} = \frac{e^{-\lambda R}\,(e^{\lambda T} - 1 + q) - q}{\lambda T e^{\lambda T}} \,. \tag{3}$$

We can call this expression $G(\lambda, q, R, T)$ or simply $G(T)$, since λ, q, and R are generally the input constants. It shows not only the average probability of a given system's being in commission, but also what fraction of a large number of such systems are in a good condition as a function of the quantities λ, q, R and T.

A simpler expression for the ready rate can be derived in the same manner, by first making the linear approximation to the exponential ($e^x \approx 1 + x$) and then proceeding as before. The process is shown in Figure 2. If the resulting equation is solved for the value of T which minimizes out-of-commission time (and thus maximizes ready rate) the result is $T = \sqrt{\frac{2qR}{\lambda}}$.

It may occur to the reader that Equation (3) can also be solved for a maximum if the exponentials are approximated linearly. Unfortunately, this gives a result which is not at all accurate, as Figure 3 clearly shows. The result is $T = v + \sqrt{v^2 + \frac{v}{\lambda}}$, where $v = \frac{qR}{1-\lambda R}$. The singularity at $\lambda R = 1$



$$A = \underbrace{\text{Area 1} + \text{Area 2}}_{= \frac{1}{2}\lambda T^2} + \text{Area 3} + \text{Area 4}$$

$$= \frac{1}{2}\lambda T^2 + (1 - \lambda T)\, qR + \lambda RT$$

$$\bar{y} = \frac{A}{T} = \frac{1}{2}\lambda T + \frac{qR}{T} - \lambda qR + \lambda R$$

$$1 - \bar{y} = 1 - \frac{1}{2}\lambda T - \frac{qR}{T} - \lambda R\,(1 - q)$$

FIGURE 2. APPROXIMATE PROFILE

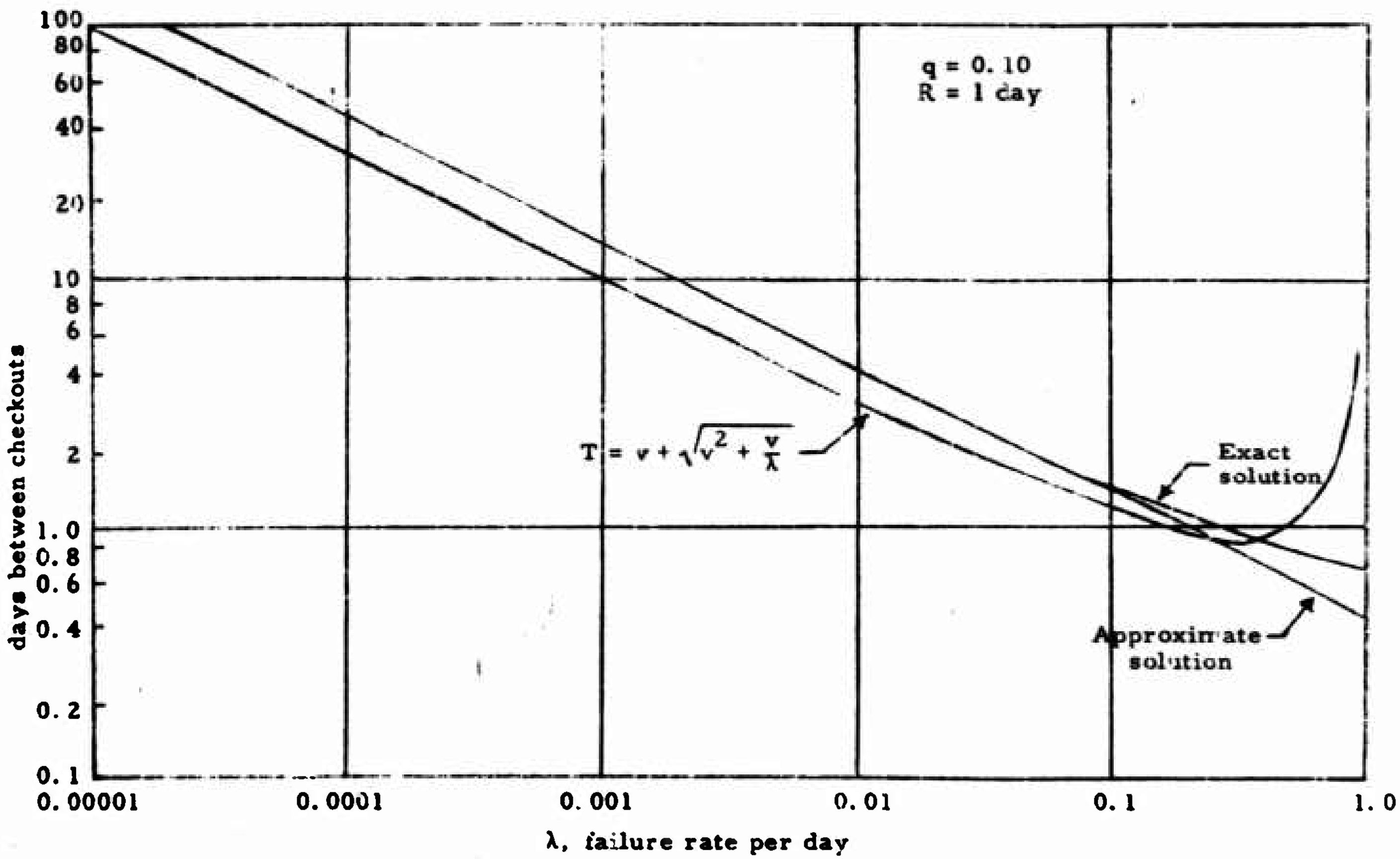


FIGURE 3

CHECKOUT INTERVAL FOR MAXIMUM READY RATE AS A FUNCTION OF FAILURE RATE (LOGARITHMIC SCALE) constitutes an additional objection. The approximate equation mentioned earlier gives by far the best solution for the maximum.

4.0 MODELS BY OTHER INVESTIGATORS

Similar analytical expressions have been developed elsewhere, one of the earlier and more prophetic[4/] ones being reported by Thompson.[(8)] He developed an equation for readiness using, for the most part, the same assumptions made here, with an important exception being that the checkout interval depends on the outcome of the immediately prior checkout. If the checkout shows a "go," or a good system, the interval is T; if "no-go," or a bad system, the interval to the next checkout is T + R. Under these conditions he showed that the ready rate is, in the symbols of this example:

$$G(T) = \frac{1 - e^{-\lambda T}}{\lambda T + \lambda R \left[1 - (1 - q)\, e^{-\lambda T}\right]} .$$

Thompson's paper also reports an equation developed at his request by Alan S. Manne, who had used the assumption of a constant interval regardless of the outcome of the checkout. In addition, Manne had used a linear approximation for the exponential, which he noted was inaccurate when the product λT exceeded approximately 0.25. His equation was identical to the one shown in Figure 2, although he used a different method to derive it.

Figure 4 shows a comparison of the ready rates predicted by the three equations when the failure rate λ is 0.10, the checkout stress q = 0.10, R = 1 day, and T varies from 1.0 to 100. It is evident that Manne's caution about the size of the product λT was very necessary. The other significant point to be observed in Figure 4 is that Thompson's equation gives a consistently higher readiness than do the equations developed here, for identical values of λ, q, and R. It does so because it does not account for "aging" or

---

4/ The basic assumptions of the Thompson Model were incorporated into an extensive "hierarchy" of readiness models developed subsequently at Space Technology Laboratories.

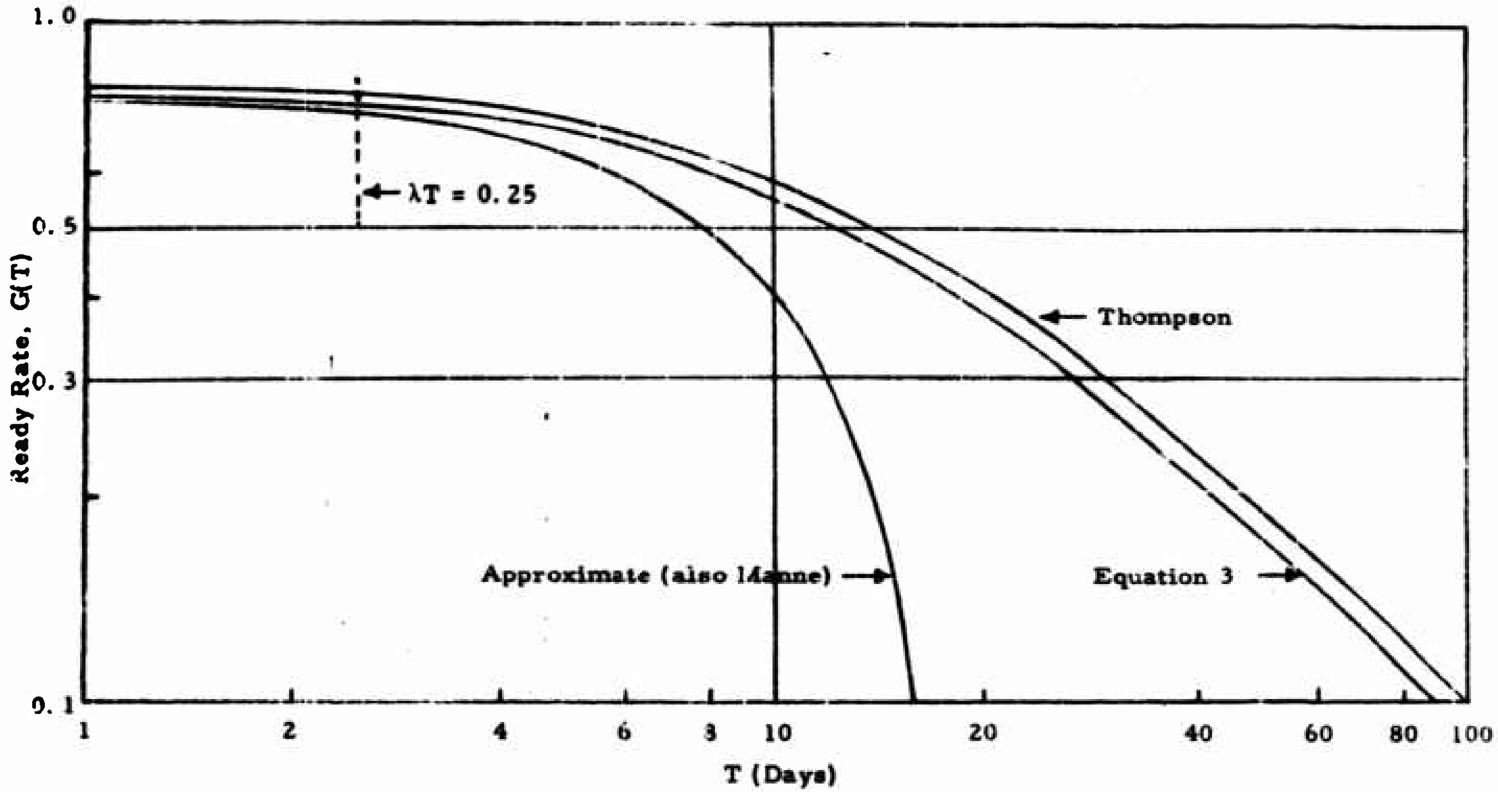


FIGURE 4

READINESS AS A FUNCTION OF TIME BETWEEN CHECKOUTS
(LOGARITHMIC SCALE)

$\lambda = 0.10$, $q = 0.10$, $R = 10$ Days failures due to standing stress during the repair of a system which failed the checkout. His derivation assumed that any repaired system was known to be good at the completion of a repair. This may be true if the entire system is replaced as a unit and is known to have no standing failure rate until installed. Generally, however, it may not be a realistic assumption, particularly when a repair consists of the replacement of only part of a system. Those parts of the system which were not defective when checked will continue to be subject to time-dependent failures during the time that a malfunctioned part is being replaced. (The case in which a system is verified by a checkout at the conclusion of a repair will be treated later.)

Figure 5 shows the readiness predicted by the same three equations when the repair time R is greatly exaggerated. As expected, the Thompson equation predicts a much higher ready rate than do the other two equations, because it neglects standing failures during the long repair periods. The differences are particularly large for very frequent checkouts. However, for the values of repair time normally encountered in practice, the differences between the Thompson equation and the one derived here are small.

In view of the discrepancies between the Thompson equation and the approximate equation, one may reasonably inquire why Thompson was interested enough in Manne's result to include it in his own discussion. The answer can be seen in both Figures 4 and 5, but it is not readily apparent. The highest point of both curves (and in fact of all three curves) occurs at approximately the same value of T. In addition, the approximate equation alone among the three can be solved explicitly for the point where its derivative vanishes, with the simple result that $T = \sqrt{\frac{2qR}{\lambda}}$; and one of Thompson's objectives was to find a simple expression for this value of the checkout interval.

Aside from the inaccuracy of the approximation used (which is usually minor), this apparently interesting checkout interval was derived without consideration of possible restrictions on support capabilities (checkout and repair) and costs. It gives a useful result only if the maximum number of systems must be kept ready, regardless of all other considerations. One such situation applies when the value of alert time on a weapon system

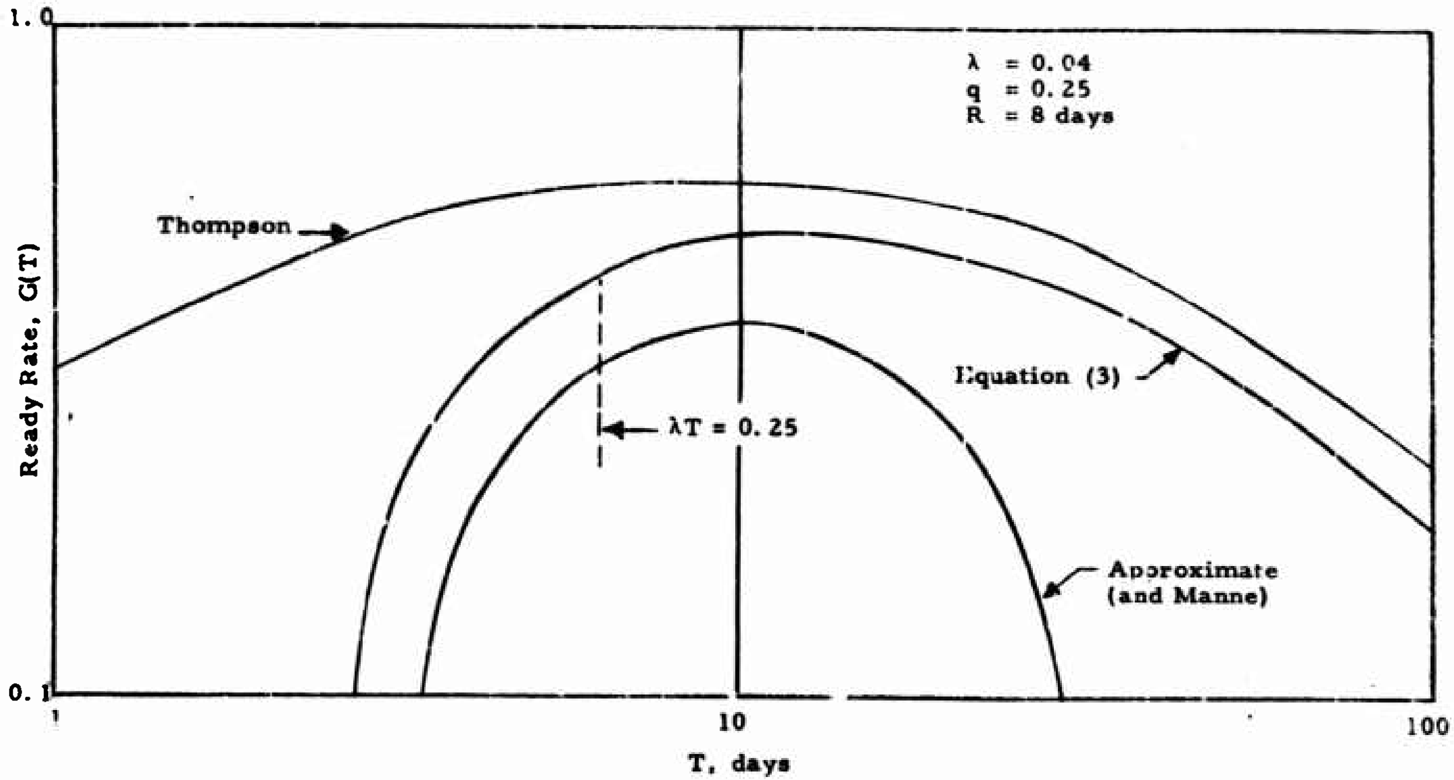


FIGURE 5

READY RATE AS A FUNCTION OF CHECKOUT INTERVAL
(LOGARITHMIC SCALE)

overwhelms all other considerations, as it does during a temporary world crisis (such as the Cuban crisis in 1962). Then, $\sqrt{\frac{2qR}{\lambda}}$ gives the checkout-interval goal for the support activity beyond which they should not try to go. Another appropriate situation would be when costs were relatively unaffected by checkout interval, as when both checkout and repair facilities are available without practical limit or added cost.

Before we leave this model, let us observe one of the less apparent limitations on its accuracy and use. It can be shown mathematically that models of this type, for determining average readiness, are exact only for an infinite time horizon. If the system is operational for a matter of years, however, and repair times are on the order of days, the error resulting from use of the equations presented is entirely negligible.

Having developed a model which may be too simple to fit many real-world cases, we can use the same technique to derive more realistic or accurate models of plausible situations.

5.0 CHECKOUT AS DOWNTIME

If we must regard the checkout process as rendering the system unusable until it is complete, we can make a relatively simple modification to the time profile to account for the fact. Figure 6 shows such a profile, with the system out of action during the checkout period of length s. The area summation is now: A = Area 1 + Area 2 + Area 3 + Area 4 + Area 5,

$$A = \int_R^{T-S} (1 - e^{-\lambda t})\,dt + (1 - q)e^{-\lambda T} \int_0^R (1 - e^{-\lambda t})\,dt + (1 - e^{-\lambda T})\,R + qRe^{-\lambda T} + s \tag{4}$$

Proceeding as before, we find that

$$G(T) = \frac{e^{\lambda(T-R)} + (1-q)(1 - e^{-\lambda R}) - e^{\lambda s}}{\lambda T\, e^{\lambda T}}$$

$$= \frac{e^{\lambda(T-R)} - (1-q)\,e^{-\lambda R} - q + (1 - e^{\lambda s})}{\lambda T\, e^{\lambda T}} = \frac{e^{-\lambda R}(e^{\lambda T} - 1 + q) - q + 1 - e^{\lambda s}}{\lambda T\, e^{\lambda T}}. \tag{5}$$

The latter form shows that the only difference between this and the previous result is that $(1 - e^{\lambda s})$ has appeared in the numerator. Since the series

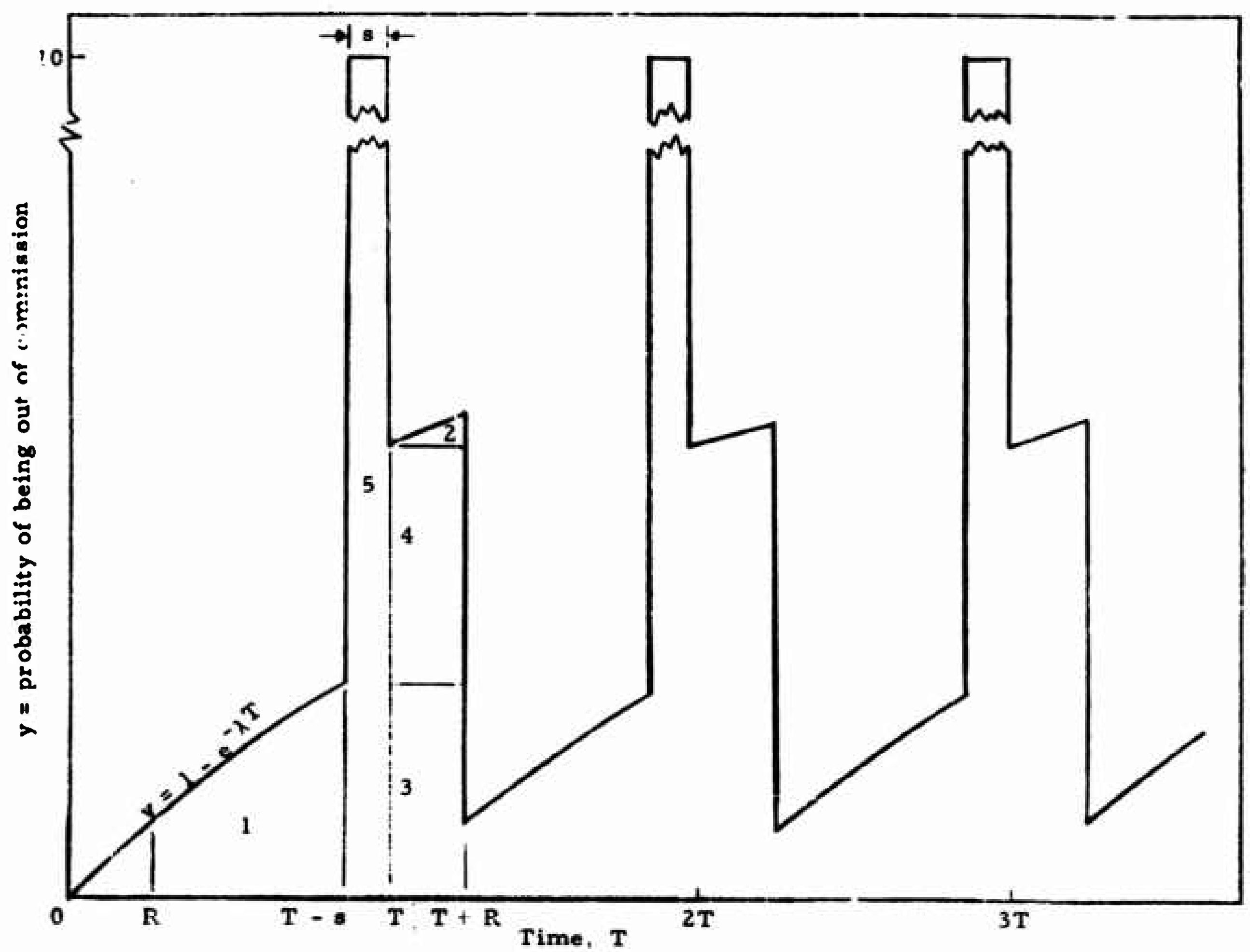


FIGURE 6. OUT-OF-COMMISSION PROFILE; SYSTEM OUT OF ACTION DURING CHECKOUT

expansion of $(1 - e^{\lambda s})$ has all negative terms $(- \lambda s - \frac{\lambda^2 s^2}{2} - \ldots)$, it can be seen that, as expected, the ready rate with checkout counted as downtime is lower than before for the same values of λ, R, q, and T. If the series approximation is made once again for the exponential (introducing the approximation before summing the areas, differentiating, etc., as in the derivation of Equation (6) and as shown by Figure 7), the result is

$$T = \sqrt{\frac{2(qR + s)}{\lambda}}$$

The checkout interval for the maximum ready rate is appreciably higher when checkout time must be considered as non-alert time.

$$A = \text{Area } 1 + \text{Area } 2' + \text{Area } 3 + \text{Area } 4 + \text{Area } 5$$

$$= \frac{1}{2} \lambda (T - S)^2 + (1 - \lambda T)\, qR + \lambda (T - S) R + S$$

$$= \frac{1}{2} \lambda T^2 - \lambda ST + \frac{1}{2} \lambda S^2 + qR - \lambda qRT + \lambda RT - \lambda SR + S$$

$$\overline{y} = \frac{A}{T} = \frac{\lambda T}{2} - \lambda S + \frac{\lambda S^2}{2T} + \frac{qR}{T} - \lambda qR + \lambda R - \frac{\lambda SR}{T} + \frac{S}{T}$$

$$\frac{d\overline{y}}{dT} = 0 = \frac{\lambda}{2} - \frac{\lambda S^2}{2T^2} - \frac{qR}{T^2} - \frac{\lambda SR}{T^2} - \frac{S}{T^2}$$

$$T = \sqrt{\frac{2(qR + S) - \lambda S\,(2R - S)}{\lambda}} \approx \sqrt{\frac{2(qR + S)}{\lambda}} \qquad (6)$$

6.0 IMPERFECT CHECKOUT; IMPERFECT REPAIR

We have been assuming that the checkout process was perfect; that is, if a defect or failure existed before a checkout or was caused by a particular checkout, it would always be discovered during that checkout.[5/] We will now account for imperfections in checkout.

[5/] Although it was not stated explicitly, it should be fairly evident that the accuracy assumption also meant that good systems would be indicated as such.

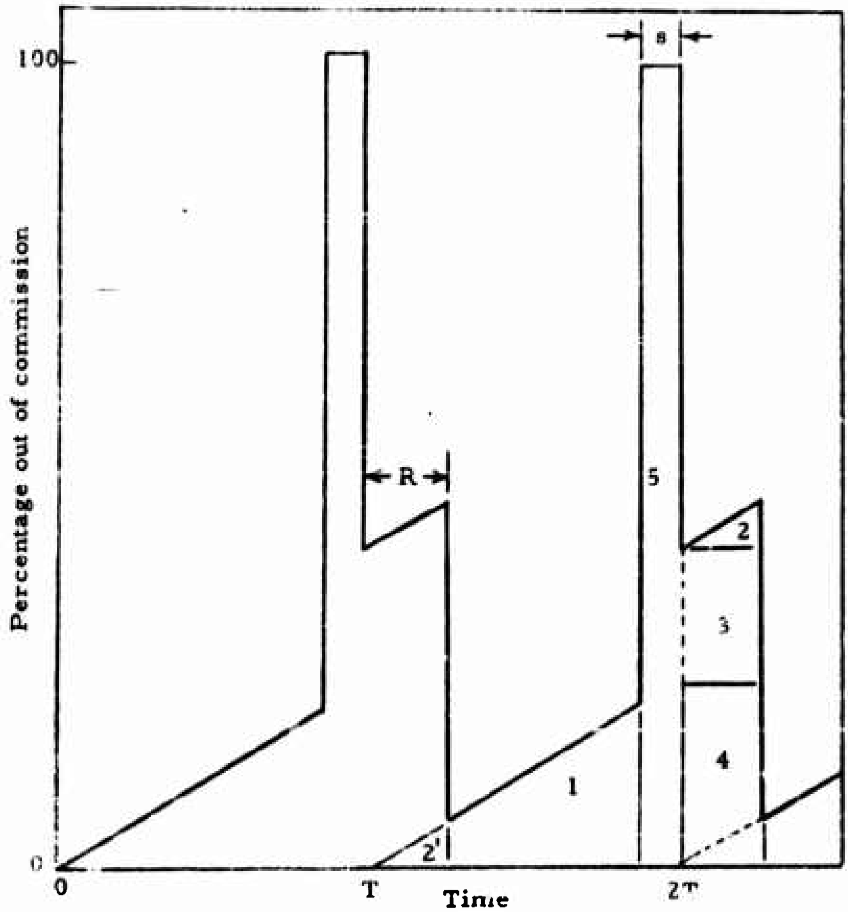


FIGURE 7. OUT-OF-COMMISSION PROFILE - APPROXIMATION (CHECKOUT AS DOWNTIME)

Two separate and distinct types of inaccuracy are possible. A checkout may call a good system bad or a bad system good. Occasions when a good system is called bad may be interpreted quite simply as cases of failure "caused" by the checkout, since such failures will ordinarily impel repair actions. In other words, such an inaccuracy (sometimes called a Type I error) can be accounted for by including in the parameter q an allowance for this type of error by substituting $q' = q + q_1$ for it, where q_1 is the probability of calling a good system bad.

Calling a bad system good is sometimes referred to as a Type II error. If such an error is random (that is, the probability of its being repeated is the same as the probability of its being made in the first place) and is equal to q_2, it can be shown that the ready rate is:[6/]

$$G = (1 - q_2)\frac{e^{-\lambda R}(e^{\lambda T} - 1 + q') - q'}{\lambda T\,[e^{\lambda T} - q_2(1 - q')]} \tag{7}$$

The checkout interval for maximum readiness can once again be found by making the linear approximation for the exponential and proceeding with the area summation, averaging, differentiation, etc.. The result is:

$$T = \sqrt{\frac{2Rq'}{\lambda\,[1 + 2q_2\,(1 - q')]}} \tag{8}$$

indicating that this time is shorter than it was for the simple method.

One other worthwhile observation can be made before we leave this topic. Consider the problem of accounting for defective repair actions. In effect, such occurrences are identical to calling bad systems good; in either case, a given fraction of the failed systems are not replaced. If we denote the probability of a bad repair as b, then simple substitution of b for q_2 in Equation (7) will give the ready rate for this case.[7/]

7.0 CHECKOUT AFTER REPAIR

We have been assuming that checkouts do not follow repair or replacement. In actuality, the reverse is more likely to be true, so we shall examine the case in which a successful checkout is required before a repair

---

6/ If the same error is repeated consistently, the ready rate goes to zero and the choice of the checkout interval is inconsequential. Such checkouts are obviously useless.

7/ The derivation of this result is fairly typical of the complications encountered in probabilistic modifications to the simple model.

is considered complete.[8/]

Figure 8 shows a portion of an out-of-commission profile under a "check after repair" policy, with the dashed line showing the "no check" situation. When the repair period ends, the profile drops only to point A' instead of A, since a probability remains that the repaired system will fail the subsequent checkout due to the probability of a bad repair (b), the checkout stress (q), or the stress of standing for R time periods ($1 - e^{-\lambda R}$). The second and third steps are progressively smaller; by the time the series dies out, the solid line may be below the dashed one. (The system being checked after repair has a shorter mean time since its last checkout and is thus less likely to have failed because of standing.) It is an extremely complicated process to derive an exact analytical expression for the area under the solid curve, because each "step" gives rise to a system of a different age when the the next checkout is due. In order to keep the fixed checkout schedule we can use the principle of superposition, separating the effects of (1) age stress and regularly scheduled checkout, and (2) checkout stress and repair deficiencies. This is an exact solution only when linear functions are involved, but the exponential is so nearly linear in the interesting range of parameters that the resulting error is ordinarily insignificant. Figure 9 shows the process in graphical form. The area due to age and regular checkout is the same as in the first derivation (from Figure 1). The area due to checkouts after repairs is:

$$R\,(\theta y + \theta^2 y + \theta^3 y + \ldots),$$

where y is the probability of a system's failing the regularly scheduled checkout$=1 - (1 - q)e^{-\lambda T}$, and θ is the probability of its failing the post-repair checkout due to faulty repair or stresses imposed during repair and

---

8/ If only a partial checkout follows the repair, to check that single function which was previously faulty, then the effect might be small enough to ignore. Another alternative might be to increase the repair time, R, to allow for checkouts and possible additional repair activity after the original repair. In general, sufficient objections can be raised to both of these approximations to make a more detailed analysis desirable.

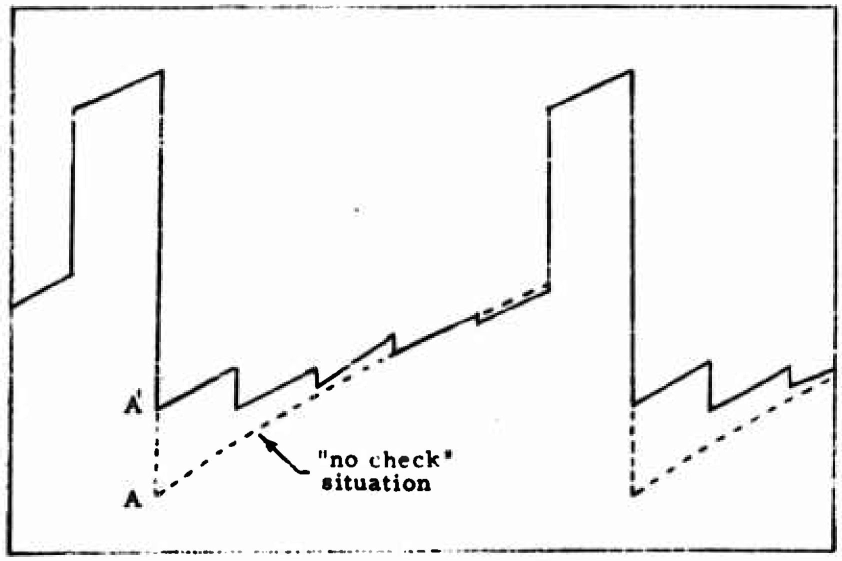


FIGURE 8

OUT-OF-COMMISSION PROFILE, "CHECK AFTER REPAIR" POLICY

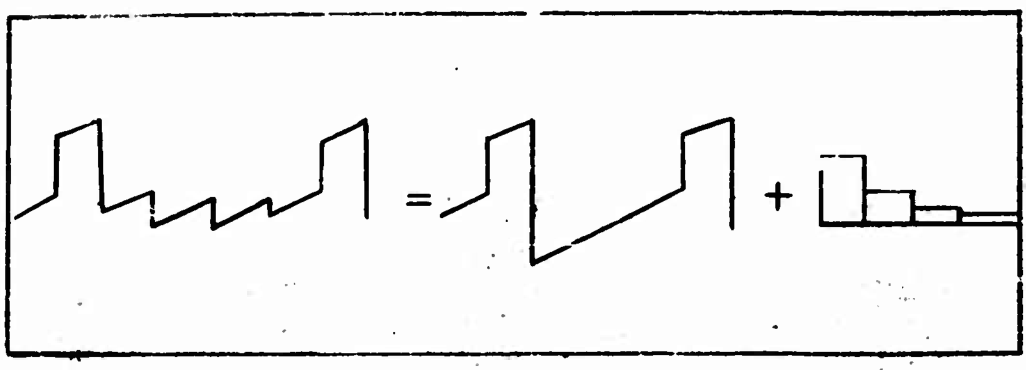


FIGURE 9

SUPERPOSITION OF NORMAL AND POST-REPAIR PROFILES subsequent checkout. (It should be evident that $\theta = b + q + 1 - e^{-\lambda R}$.) Since the series is a binomial, the area involved can be expressed as:

$$R\theta\left[\frac{1 - (1 - q)\, e^{-\lambda T}}{1 - \theta}\right]$$

Combining areas, dividing by cycle length, and subtracting from one, the ready rate can be found to be:

$$G(T) = \frac{(e^{\lambda T} - 1 + q)\,(e^{-\lambda R} - \frac{\lambda R \theta}{1 - \theta}) - q}{\lambda T\, e^{\lambda T}} \tag{9}$$

Once again, the approximate checkout interval for maximum readiness can be found by making the linear approximation to the exponential, the result being:

$$T = \sqrt{\frac{2qR}{\lambda(1 - \theta)}}\ . \tag{10}$$

Again, the checkout interval which gives the maximum readiness is lengthened from what it was in the original simple model. It should be clear at this point that modifications designed to improve the validity of the simple model may increase or decrease the interval at which the maximum ready rate occurs.

Since we now have expressions which show the ready rate with and without checkout after repair, we can use them to compare the two policies and find which one gives a higher ready rate. A comparison in parametric form of Equations (7) and (9) gives no simple result, so that an evaluation of G for both cases must be made.

It should be evident that the peculiarities of a particular application can be incorporated into a model relating readiness to other parameters, and particularly to the checkout interval. For the cases discussed here, the final approximate result for the checkout interval which maximizes readiness is:

$$T = \sqrt{\frac{2\,(R\theta + q)}{\lambda\,[1 + 2q_2\,(1 - q)]\,(1 - \theta)}}\ . \tag{11}$$

In general, the effects of the complicating factors studied here are small compared with the effects of the parameters λ, q, and R.

8.0 THE EFFECT OF MAINTENANCE RESOURCES

Another source of weapon system unavailability arises when a failed system must wait for maintenance resources (men and equipment) because they are engaged in repairing another system. Since demands for such resources are typically distributed randomly along a time scale, closely-spaced demands (i. e., failures requiring service) will occasionally outstrip the repair capability unless support is provided at an unrealistically (and uneconomically) high level. The amount of unavailability which results from this source can in many cases be legitimately and accurately treated by the branch of mathematical science called queuing theory, or waiting-line theory.

The basic information requirements are estimates or factual data concerning the frequency of demands for a particular resource and the average length of time that resource will be occupied in fulfilling the demands. (If economics enters the evaluation then the costs of providing resource increments and the value of the resulting incremental operational capability must also be known.) As an example, the Minuteman weapon system specifications provide such workload factors.[9-12] Failure rates for systems and subsystems are estimated, in terms of failures per month for a 150 missile wing. "Time lines" indicate how long each type of resource will be tied up in performing each task, including travel times to site, etc.. The product of these two pieces of information, for a given equipment item or a particular maintenance team, will yield the expected monthly utilization in hours.

If one is concerned only with resource utilization, the necessary quantity can be determined easily and directly. Suppose, for instance, that a particular major component has an estimated failure rate of 50 per month for a 150 missile wing, and that each failure can be expected to tie up a certain piece of equipment and associated team for 50 hours. Expected utilization is then 50 x 50 or 2500 hours per month. Assume 250 hours per month permissible utilization of the equipment (perhaps 8 hours per day, 31 days a month), and it is seen that 10 sets of in-commission equipment are required with this method of calculation. Assume 140 hours per month per team, to allow for leave, illness, squadron duties, etc., and 18 teams of maintenance personnel need to be assigned. Utilization of teams and equipment alike would be nearly 100 per cent, a very "efficient" operation. A substantial part of the force would be kept waiting its turn in line, however, unless each new failure occurred just as the previous one had been corrected -- and we know that this does not happen. There are those days when all goes well, and those when everything goes wrong.

Large bodies of data indicate that complex weapon systems usually experience malfunctions distributed in a "random" fashion. For a specific example, let us assume data has shown or engineering estimates indicated that 3 per cent of the 150 missile wings will malfunction on an average day; thus there will be an expected 135 failures per 30-day month. For the moment also assume there is one particular type of equipment and maintenance team which handles these failures, and that each failure takes exactly 2 days to repair. Using a table of random numbers(13) and applying the 3-per-cent failure probability, the following failures occured in the simulation of a 30-day month.9/

9/ With a 3-per-cent failure probability for each of 150 missiles, the expected number of failures per day is simply 3 per cent of 150 or 4.5. With this rate, the probability of having a particular number of failures on any given day is, from a table of Poisson's Exponential Binomial Limit:(14)

No.	Probability	No.	Probability	No.	Probability
0	0.011	5	0.171	10	0.010
1	0.050	6	0.128	11	0.004
2	0.112	7	0.082	12	0.002
3	0.169	8	0.046	13	0.001
4	0.190	9	0.023	14	0.000

We now draw three-digit random numbers. If the number drawn is from 1 to 11, there are no failures that day; if from 12 to 61, one failure; if from 62 to 173, two failures, etc..

Day	No. of Failures	Day	No. of Failures
1	8	16	2
2	11	17	4
3	4	18	1
4	6	19	4
5	5	20	3
6	3	21	0
7	3	22	3
8	4	23	5
9	1	24	7
10	3	25	4
11	0	26	4
12	1	27	5
13	4	28	3
14	3	29	4
15	1	30	4

This month has only 110 failures, instead of the expected 135. In some months there will be substantially more failures than the average expected. Figure 10 shows this months distribution. There were 2 days with no failures, 9 days with 4 failures, 1 day with 11, etc.. The "expected" distribution is also shown. Despite the lower-than-usual monthly workload, we can expect some workload control problems, especially since the two heaviest days fall one after the other, on the 1st and 2nd of the month. Figure 11 indicates the workload for the month, assuming that each job is discovered and started at the beginning of one day, and is completed at the end of the second day. A carryover of 4 jobs from last month (assumed, based on last day this month) plus 8 more on day 1, occupies 12 teams; the carryover of 8, plus 11 on day 2, results in the peak load of 19 teams. No new work on day 11 and only 1 job on day 12, results in almost no work by contrast.

To perform this month's work without any delays whatever will require 19 teams on duty, although team utilization would only be at 39 per cent. Suppose we had only 18 teams; the unit of work numbered (1) would have to be delayed from day 2 to day 3. With only 14 teams, units of work numbered (1) through (6) would have to be delayed for the periods indicated. There would be a total of 13 alert missile days lost. Table I indicates the story as teams are cut progressively to 9, where 172 alert missile days are lost waiting for teams (nearly 6 missiles on the average each day), and yet team

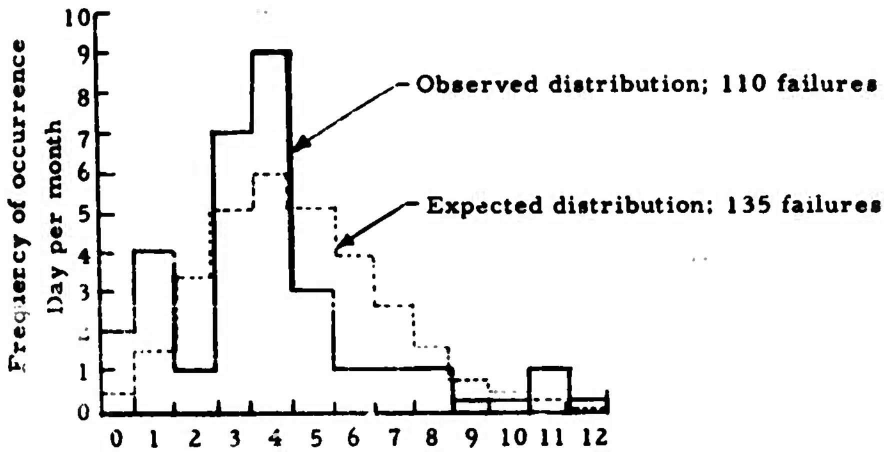


FIGURE 10

DISTRIBUTION OF FAILURES FOR A SIMULATED MONTH

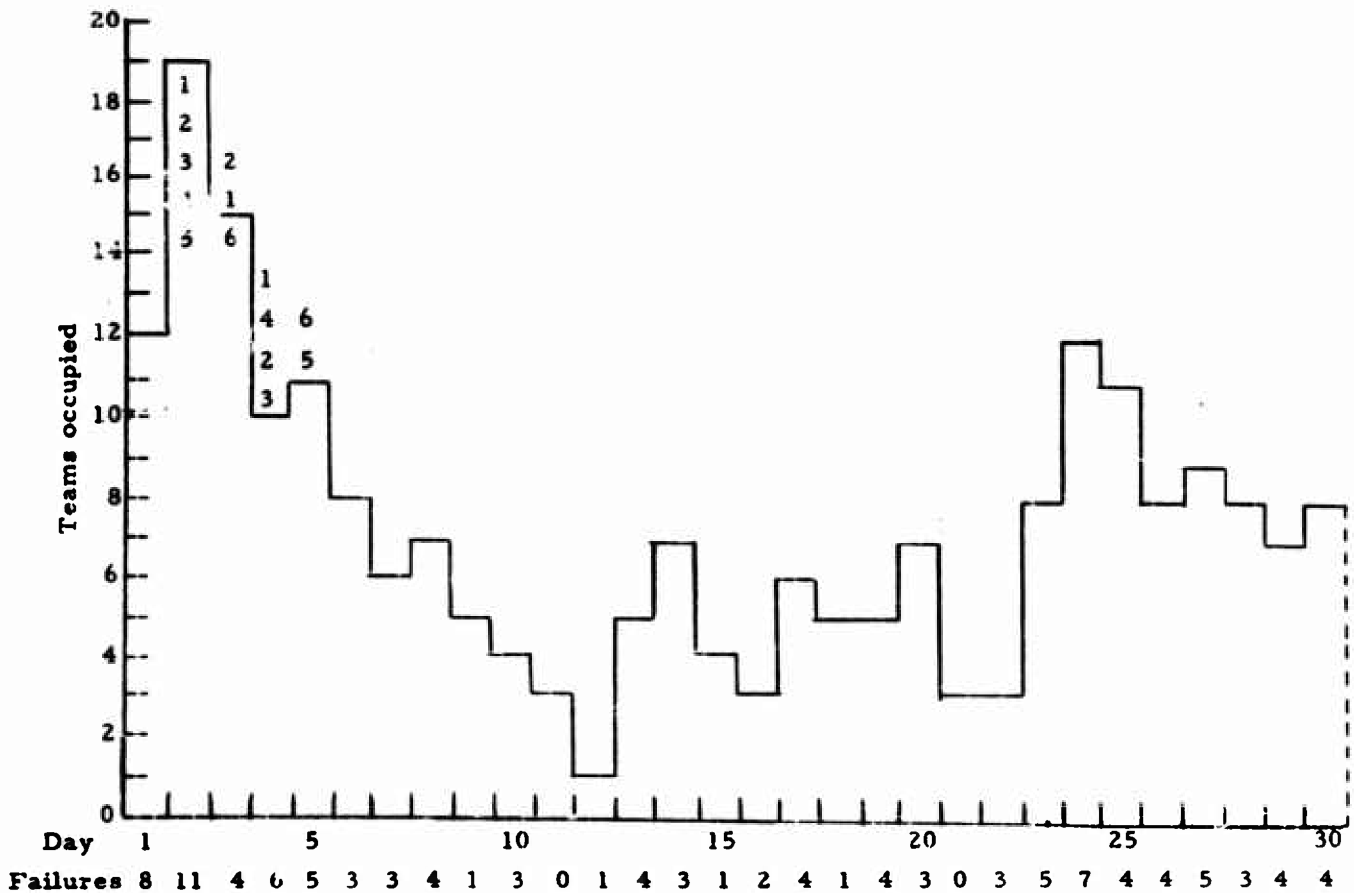


FIGURE 11

WORKLOAD DISTRIBUTION FOR A SIMULATED MONTH utilization is only 82 per cent.

TABLE I

MISSILE WORK DELAYS AS A FUNCTION OF NUMBER OF TEAMS FOR A SIMULATED MONTH

No. of Teams	Percent Utiliza-tion	Alert Missile-Days Lost Wait-ing for Teams
19	39	0
18	41	1
17	43	2
16	46	5
15	49	8
14	52	13
13	57	22
12	61	35
11	67	62
10	73	105
9	82	172

We have already indicated that this month's experience was "unusual," both in having fewer failures than expected and in the failure distribution. We will enter the table of random numbers again, therefore, to establish a second month's experience. Figure 12 shows the distribution, and Figure 13 the workload. This time there were 131 failures (vs. the 135 expected). Table II presents the story as the number of teams are reduced, as did Table I for the first month.

TABLE II

MISSILE WORK DELAYS AS A FUNCTION OF NUMBER OF TEAMS FOR ANOTHER SIMULATED MONTH

No. of Teams	Percent Utiliza-tion	Alert Missile-Days Lost Wait-ing for Teams
14	62	0
13	68	1
12	73	7
11	80	18
10	88	42
9	97	132

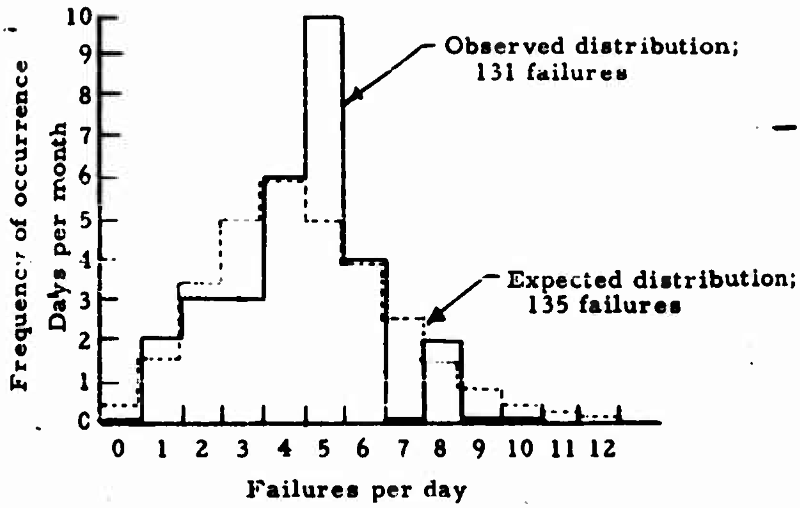


FIGURE 12

DISTRIBUTION OF FAILURES FOR A SECOND MONTH

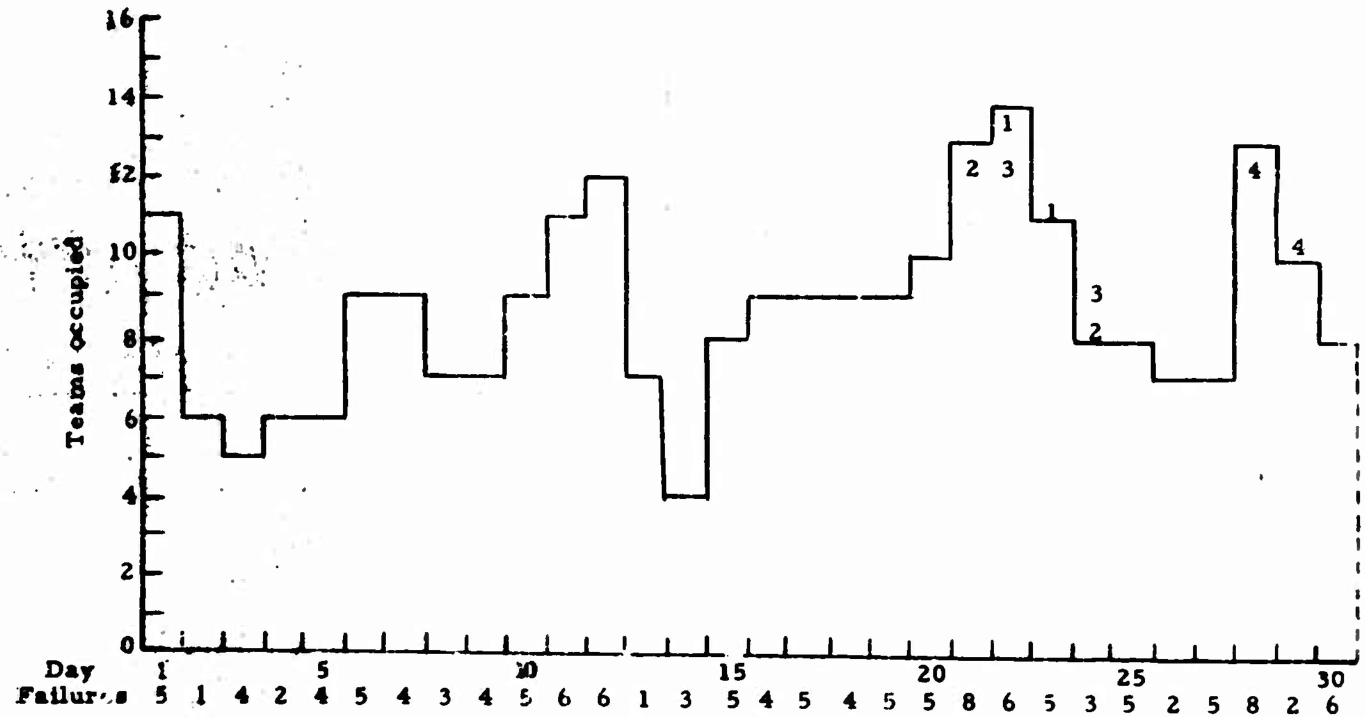


FIGURE 13

WORKLOAD DISTRIBUTION FOR A SECOND MONTH

Again the incompatibility of high utilization of maintenance resources and a high degree of operational readiness is evident. At the extremes of the two tables, we choose between team utilization of only 40 to 60 per cent and maximum possible force readiness, or team utilization of 80 to 95 per cent with an average five or so missiles waiting for maintenance at all times.

In 1958, Peck and Hazelwood produced a set of finite queuing tables (Reference 15) in which they stated:

> This monograph is intended to provide useful tables for the solution of a variety of queuing problems. There are several textbooks and monographs which discuss the theoretical aspects of queuing theory, and the literature contains an impressive number of research papers dealing with such problems. However, there is a paucity of directions for useful applications of these ideas. A person exposed to the concepts of queues is able to recognize broad areas where the theory could be applied, but he is often at a loss to find ways of solving the problems. ....
>
> In general, queuing theory deals with the formation of a queue, or waiting line. Suppose there is some service point, such as ... a repairman servicing a broken machine, etc. Part of the time this service point will be busy providing service, part of the time it will be idle. If the service point or "channel" is busy and another customer arrives, he must wait. This forms a queue.

This situation fits our missile force where missiles or aerospace ground equipment and facilities, requiring maintenance, are the "customers," and maintenance teams are the service point. Using our hypothetical example presented earlier, let us enter the Peck and Hazelwood tables and examing the findings. A typical portion looks like the following:

POPULATION, N = 150

X	M	D	F
.025	7	.096	.999
	6	.215	.998
	5	.438	.992
	4	.785	.966
.026	7	.113	.999
	6	.247	.997
	5	.487	.990
	4	.838	.955
.028	7	.154	.999

The population represents the number of "customers" which can demand service -- in the Minuteman case 150 missiles. X = service factor = $\frac{T}{T + U}$ where T = average service time or repair time, and U = average time not calling for service 10/ -- in our case the mean time between failures. M = service channels or number of repair teams. D = probability that if a unit calls for service it will have to wait ("delay" probability). F = efficiency factor = $\frac{H + J}{H + J + L}$ where H = average number of units being serviced, J = average number of units running (on alert), and L = average number of units waiting for service.

Assumptions in our specific example were that 3 per cent of the wing would malfunction each day, on the average, and take 2 days to repair. Thus, T = 2, U = 33.3 days, and X = $\frac{2}{2 + 33.3}$ = .057. The tables have X = .056 and X = .058. Rather than interpolate, the section for population 150, X = .058, is reproduced below.

X	M	D	F
.058	14	.060	.999
	12	.198	.997
	11	.330	.993
	10	.518	.985
	9	.747	.962
	8	.939	.905

L, the average number of units waiting for service, can be represented as L = N(1 - F). Thus with 9 teams, for instance, L = 150(1 - .962) = 5.7 missiles. This is 5.7 x 30 = 171 missile days per month.

Table III compares the queued missiles vs. teams for the random month of Tables I and II respectively, and per Peck and Hazelwood. Notice that P & H values are compatible with those of the two random months. We re-emphasize here that one random month is wholly inadequate to properly determine the distribution of demands, these examples being given only to illustrate the principles.

We hope that two major points are clearer now: (1) high utilization of resources in the correction of critical malfunctions (those preventing launch of the missile) is incompatible with high mission readiness, and therefore uneconomical; (2) it should be possible to develop a technique that can use

---

10/ Note that for the elementary problem involving only two states (either "GOOD" or "IN REPAIR"), service factor is simply 1 - availability.

TABLE III

MISSILE WORK DELAYS VS. NUMBER OF TEAMS
Peck and Hazelwood; Tables I & II

No. of Teams	Missile Days Queued Table I	Table II	P & H
14	13	0	4.5
12	35	7	13.5
11	62	18	31.5
10	105	42	67.5
9	172	132	171.0

cost-effectiveness criteria to measure the interaction between number of teams available and number of missiles or related equipment waiting for service.

9.0 THE ESTIMATION OF PARAMETERS

RAND/AFLC/SAC Study: In December 1962, the Logistics Department of the RAND Corporation undertook a project, in collaboration with Headquarters AFLC, to implement the results and methodology of RM-2578, "Determining Checkout Intervals for Systems Subject to Random Failures." The system chosen for the initial application was the Atlas "D" weapon system, which was felt to have been operational long enough to generate some reliability and time-line experience, both of which were essential inputs to any implementation action. During early February 1963, assistance was requested by the Directorate of Requirements, Headquarters SAC, "to determine optimum time intervals between operational exercises for Strategic Air Command Intercontinental Ballistic Missiles." 11/ At this point, it was decided to combine the two efforts, concentrating primarily on the Atlas "D".

11/ Reference unclassified letter from Col. R. E. Barton, Directorate of Operations, Hq. SAC to The RAND Corp., dated 5 Feb. 1963; subject: Determination of Optimum Intervals between ICBM Weapon Systems Operational Exercises.

First, the objectives of the study program will be discussed, followed by discussion of a machine simulation that generates the required information to plan a test program, and also demonstrates how the secondary objectives can be achieved.

Objectives: The objective of the AFLC portion of the study was broadly comprehensive, but can be simply stated: to develop a methodology for setting maintenance activity intervals on normally inert systems. The SAC objectives were explicitly stated in their request for assistance. The primary objective was to determine optimum time intervals between operational exercises for SAC ICBM's -- "optimum" meaning the interval that would yield the most ready weapons. The secondary objective was to be able to estimate the lift-off capability of the ICBM force at any arbitrary point in time, that is, how many will be launched. A third objective was to derive the confidence level associated with the estimate of the lift-off capability.

Assumptions[12/]: Certain assumptions are applicable to this study effort, as follows:

(1) System failures, when not in checkout, are Poisson distributed.
(2) The probability of system failure in checkout may be represented by a constant.
(3) The probability that a failed system will be assigned to alert status is zero.
(4) Type I and Type II statistical errors are negligible.

Data Requirements: In order to estimate q and λ for the related mathematical models, four inputs are required. Two of these are reliability inputs; two are experience-time inputs. The first reliability input is the decay rate while standing; that is, the rate at which defects occur while the missile is on strategic alert. This can also be expressed as a mean time

12/ For a partial relaxation of these assumptions, see Volume III, Example B of TG-II report.

between failures (MTBF). The second input is also a reliability parameter; it is an index of the stress of an exercise, perhaps more properly thought of as the probability that an exercise will trigger a failure in a system that was otherwise good. The time inputs needed are the time off alert for the exercise itself, and the time off alert to correct a failure, whether it was caused by the exercise or was a standing failure uncovered by the exercise.

Test Conditions: The decay rate and the stress of an exercise are both reliability parameters and are ordinarily observed together with no practical way of separating them unless a statistical experiment is conducted.[13/] Figure 14 is a diagram of the type of statistical experiment that can be used to separate the effect of time failures and the effect of exercise-caused failures. It consists of some sort of starting transient (a successful exercise is a good beginning) followed by back-to-back exercises at intervals t. This is a cyclic procedure that is repeated until a satisfactory number of experiences have been recorded, at which time the test phase is discontinued, proceeding to an operational phase shown on the same figure; namely, individual countdowns held at intervals T, which intervals are to be determined from the results of the previous tests, according to the rules in RM-2578. The reason back-to-back exercises are suggested is to permit separation of the two reliability effects, namely, the standing failure rate and the probability of an exercise causing a failure. Note that the first of the two back-to-back tests in each cycle is subject to the time failures built up during the period of length t plus the stresses associated with the exercise itself. The second of the two exercises, however, takes place after only a very short time has elapsed and has almost no chance to build up time failures. It is subject, for all practical purposes, only to the stresses of the exercise itself. Experience with the second exercises provides an estimate of the stress of the exercises. Subtracting this effect from the experience with the first exercises provides a measure

---

13/ For some types of malfunctions, detailed failure analysis may permit an estimate of the relative importance of time and exercises. This is a severely limited capability.

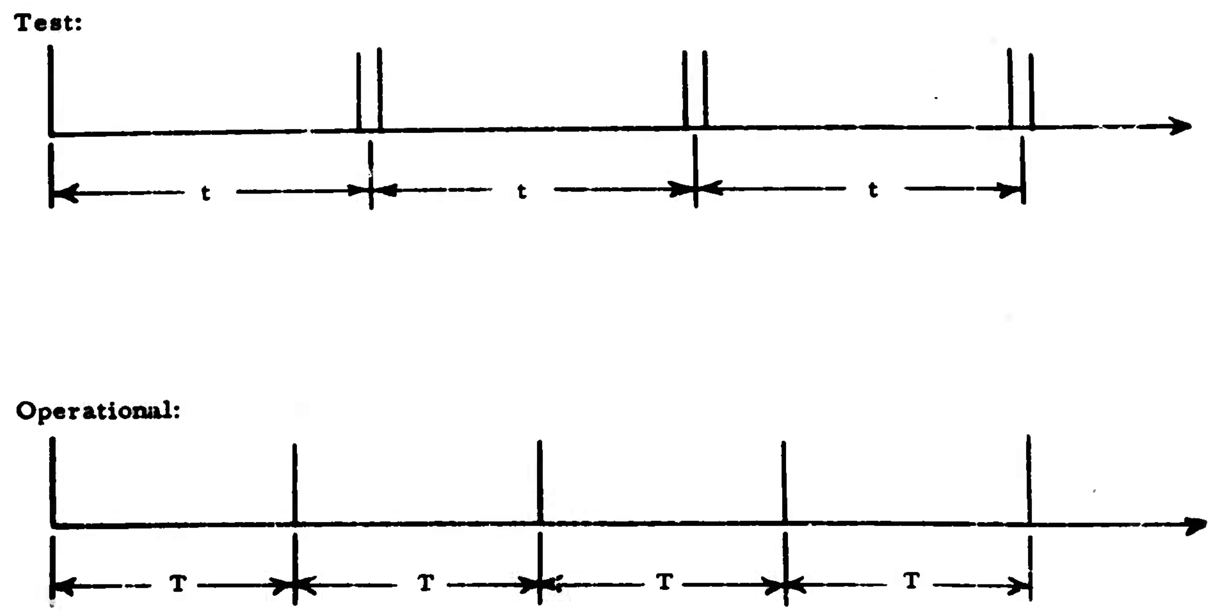


FIGURE 14

DIAGRAM TO SEPARATE EFFECT OF TIME FAILURES AND EFFECT OF EXERCISE-CAUSED FAILURES of what portion of the defects in those exercises were due to time, and time rate of decay can be computed. Thus, it is assumed that the model of reliability (the probability of no defect during the exercise) is $(1 - q)e^{-\lambda t}$, where q is the probability that an exercise will trigger a failure in an otherwise good system, λ is the failure rate while the system is on alert, and t is the time since the system was last verified. Thus, the second exercise in the back-to-back cycle is used to determine q, and the first exercise when combined with this information gives the estimate of the decay rate λ.

To see how the process works in more detail, beginning with the expression for successful exercise:

$$\text{Probability of no defect} = (1 - q)e^{-\lambda t}, \text{ and}$$

taking natural (Naperian or Base e) logarithms of both sides, leads to:

$$\log\ [\text{Probability of no defect}] = \log\ (1 - q) - \lambda t.$$

Since both the probability of no defect and the quantity $(1 - q)$ are by definition less than 1.0, the logarithms of these quantities are (by definition of the logarithm) negative numbers. Thus, each term of the expression is in fact (if not in appearance) negative, and this expression, which has the classical linear form $y = a + bx$ can be plotted, as in Figure 15.

Comparing this with the previous equation $(1 - q)e^{-\lambda t}$ above, it is seen that the intercept is simply the negative logarithm of $(1 - q)$, and the slope is simply the decay rate λ. Thus by plotting a diagram like Figure 15, the relation between the outcome of exercises at zero time and time t, and the desired parameters is seen graphically. Note that exercises conducted at any two different intervals would generate estimates of the parameters. The selection of zero (i.e., back-to-back) as one of these times maximizes the efficiency of the procedure, from the standpoint both of effort required and accuracy of results. 14/

14/ J. Kiefer and J. Wolfowitz, "Optimum Designs in Regression Problems," Annals of Mathematical Statistics, Vol. 30, 1959, pp. 271-294.

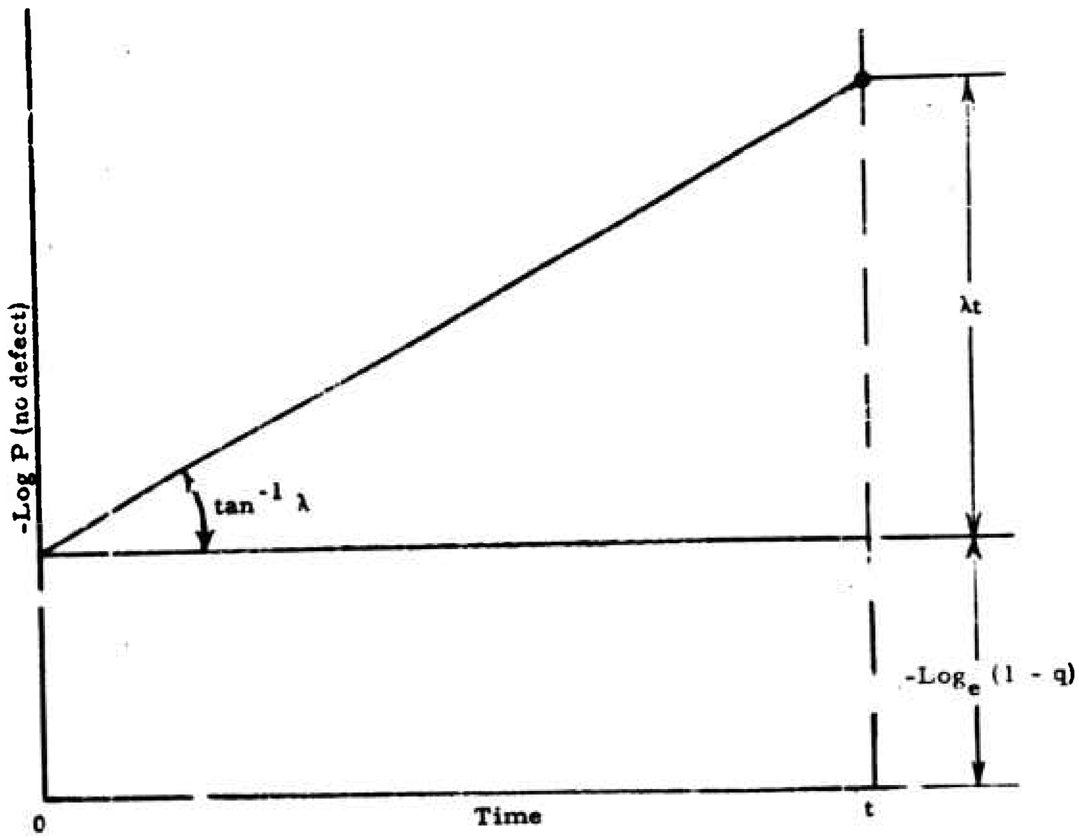


FIGURE 15

DIAGRAM OF THE RELATION BETWEEN THE OUTCOME OF EXERCISES AT ZERO TIME AND TIME t

Optimum Test Conditions: As far as the test phase of weapon system operation is concerned, the question to be answered is, what should be the value of t to get the most information out of testing. The problem will be examined from the standpoint of a statistically best value of t. (A problem which has been studied separately is the value of t which provides the best alert capability during the test phase, and the value of t which is compatible with the support capability assigned to the Atlas squadrons in the field.) The best frequency or interval between back-to-back exercises, as far as the statistical problem is concerned, could be worked out analytically; but because of concern for certain traps it is preferable to do things a little differently. The procedure is as follows: first, set arbitrary reliability values on the nine missiles in an Atlas "D" squadron. In other words, values of λ and q are set arbitrarily, and values of time off alert for exercising and time to make a repair if one were required are used. These latter values were obtained from Air Force personnel at Vandenberg on the basis of prior experience with the Atlas "D" system. The second step is to simulate the exercise schedule, namely to simulate the performance of back-to-back tests in an Atlas "D" squadron. In effect, this is analogous (using a 7090 computer) to throwing dice to find out whether a given missile passed an exercise or failed it. The outcome is then analyzed, that is, estimates of the value of the decay rate, λ, the stress of the exercise, q, and the lift-off capability of the squadron, z, are computed.[15] These computed values are then compared with the real values which had been set for these particular missiles in this particular run. These steps were repeated 999 times to generate an experience sample; that is, to see what would happen in the world of probability if such an experiment were to be run on a system which had particular reliability characteristics. Results were then analyzed to see how good the estimates were on the average and how bad they might prove in some unusual cases. A wide range of reliability parameters were covered. For example, a range of λ was covered which

---

15/ Using Eq. 2.5 (p. 18) of RM-2578 to obtain the ready fraction, and mutiplying by (1 - q) to determine the fraction successfully counted down.

was indicative of mean times between failure all the way from 10 days to 320 days; values of q were covered all the way from 0.1 through 0.5. The generalized results are shown in Figure 16 for the case of a calendar time limit, specifically 180 days. [16] In all cases the standard error of prediction, sometimes called the standard deviation, could be reduced by decreasing t to about 36 days. This is true regardless of the value of q (three values being shown here) and regardless of the value of mean time to failure. Thus, with a limitation on calendar time, the statistically optimum value of t for the test phase is approximately 36 days. A by-product of the simulation program is the generation of a confidence map (Figure 17) which shows the confidence interval which can be associated with the prediction of the lift-off capability of a squadron. If the lift-off capability predicted by the tests is entered at the bottom of the chart, the width of the confidence interval is then read from the lines shown for 50, 80, or 90 per cent confidence. The dashed line shown across the chart is indicative of how it was constructed; a missile with a particular known reliability (or lift-off capability) was inserted into the simulation, and the results for 50, 80, and 90 per cent confidence were plotted for this missile and several others with different reliabilities to form this chart. At this point, the three objectives stated in the SAC letter of 5 February have been achieved: to determine the best exercise interval for the ICBM, to get a measure of the lift-off capability, and to find the confidence intervals associated with that measure of capability.

10.0 CONCLUSIONS

In this Example a methodology has been developed for establishing, by test, data necessary for determination of appropriate intervals for performing scheduled maintenance actions on ICBM's. It has been shown through a simulation model how these data are developed, and how both

---

[16] If instead of calendar time, the total number of exercises is limited, then it can be shown that t should be approximately 1.5 times the MTBF. H. Chernoff presented a similar result in "Locally Optimal Designs for Estimating Parameters," Annals of Mathematical Statistics, Vol. 24, No. 4, 1953, p. 586.

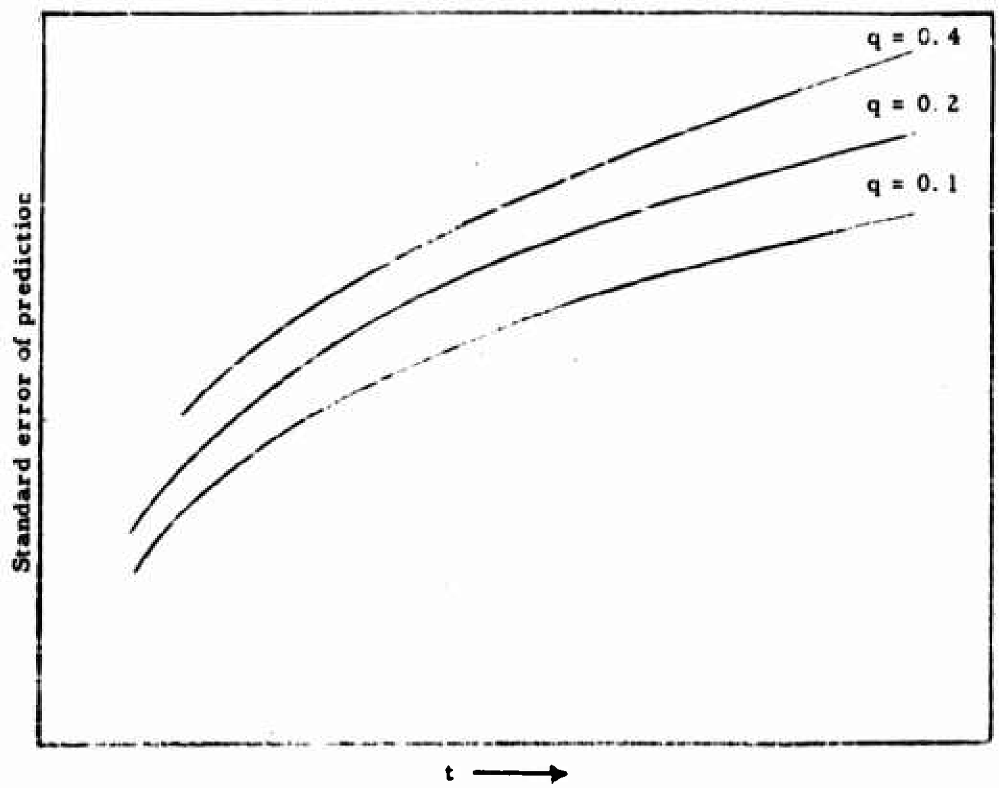


FIGURE 16

GENERALIZED RESULTS

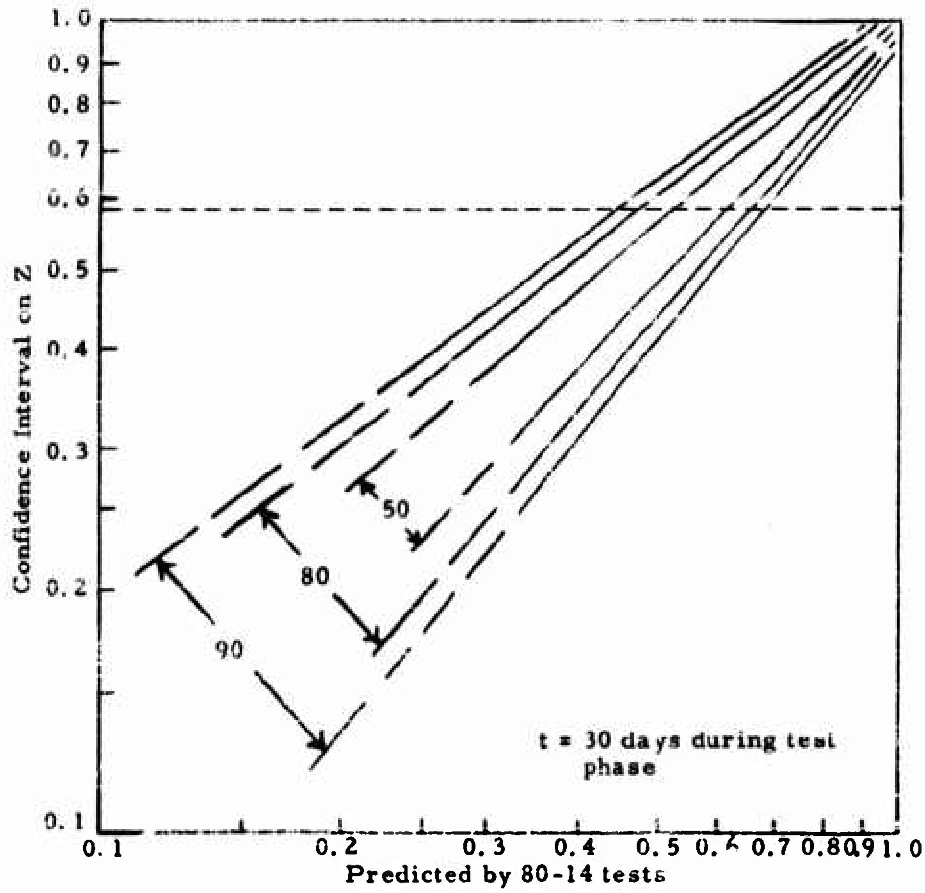


FIGURE 17

CONFIDENCE MAP

maintenance intervals and operational capability estimates are determined from them.

Finally, it is concluded that, unless the critical data elements of exercise stress, decay rate, exercise time, and repair time are all known, separately and distinctly, operational capability cannot be predicted nor can maintenance policy be validated.

REFERENCES

(EXAMPLE E)

1. Carhart, R. R., "A Survey of the Current Status of the Reliability Problem," RM-1131, AD 80637, The RAND Corp., August 14, 1953.

2. Davis, D. J., "An Analysis of Some Failure Data," Journal of the American Statistical Association, Vol. 47, No. 258, pp. 113-150, June 1952. (Or see idem, P-183, The RAND Corp., February 12, 1952.

3. Boodman, David M., "The Reliability of Airborne Radar Equipment," Journal of the Operations Research Society of America, Vol. 1, No. 2, pp. 39-45, February 1953.

4. Lusser, Robert, "Electronics, The Key Problem of Reliability," Proceedings of Joint Military-Industry Guided Missile Reliability Symposium, Redstone Arsenal, Huntsville, Alabama, October 15-17, 1956.

5. Kamins, M., "Validating Maintenance Policies and Estimating Launch Capability for Ballistic Missiles (U)," RM-3645, The RAND Corp., (Secret)

5. Cox, D. R., and Smith, W. L., "On the Superposition of Renewal Processes," Biometrika, Vol. 41, pp. 91-99, 1954.

6. Morse, Philip M., Queues, Inventories and Maintenance, p. 165, John Wiley and Sons, New York, 1958.

7. Thompson, William B., "Employment of Launch Site Test Equipment for Maximum System Reliability," Technical Military Planning Operation. SP-71, General Electric Co., Santa Barbara, California, February 1. 1960.

8. Ballistic Systems Division, "Qualitative Personnel Requirements Information for WS-133A Minuteman Hardened and Dispersed (U)," Vol. 1, D2-5859, July 1962, (Secret).

10. Ballistic Systems Division, "Maintenance Ground Equipment Failure Rate Data (U)," Vol. II b, D2-6951, revised April 1962, (Secret).

11. Ballistic Systems Division, "Minuteman Operational Support Planning Failure Rates (U)," BSO-169-1, May 15, 1962, (Secret).

12. Knowles, H. F., "MTBF Apportionments for Production Equipment (U)," EM-7667-2, Autonetics, May 21, 1962, (Secret).

13 The RAND Corp., A Million Random Digits with 100,000 Normal Deviates, The Free Press, Glencoe, Illinois, 1955.

14. Molina, E. C., Poisson's Exponential Binomial Limit, D. Van Nostrand Co., Inc., New York, 1957.

15. Peck, L. G. and Hazelwood, R. N., Finite Queuing Tables, John Wiley and Sons, Inc., New York, 1958.

EXAMPLE F

A VULNERABILITY MODEL FOR WEAPON SITES WITH INTERDEPENDENT ELEMENTS [1]

1/ This example is taken from "A VULNERABILITY MODEL FOR WEAPON SITES WITH INTERDEPENDENT ELEMENTS," RAND P-1384, S. I. Firstman, May 27, 1958; also published in Journal of ORSA, pp 217-225, March-April 1959.

ABSTRACT

This example describes a simple "counting" model, employing probability grid transparent overlays, which aids in the determination of the trade-offs, measured in survival probability, between site dispersal and hardening for a weapon complex composed of several interdependent elements, separated by distances of less than two lethal radii. The survival-probability expressions are obtained through the use of Markov chains. An example of vulnerability estimation is given.

TABLE OF CONTENTS

ILLUSTRATIONS

TABLES

1.0 INTRODUCTION

The model to be described was developed in order to determine the tradeoff, measured in survival probability, between site dispersal and hardening for a weapon complex composed of several interdependent elements which are separated by distances of less than two lethal radii. This simple "counting" model employs overlays, probability grids, and Markov chains. It is useful in assessing damage effects when the complex is subjected to enemy attacks of varying size, yield, and accuracy. The complex could, for example, be a missile base, consisting of several launchers and a command and guidance site.

Broadly, the problem considered is that of determining the vulnerability due to enemy action of a weapon-complex composed of several interdependent elements (e.g., missiles and a guidance and command station), where the elements are separated by distance of less than two lethal radii, as determined by the site hardness and attacking-weapon yield. [2/] Conversely, the probability of survival of exactly 0, 1, 2, 3, . . . weapon elements and the associated command-unit(s) is determined for an enemy attack of a specified magnitude, against the complex. In order for each weapon-element to be used after an attack, both it and an element called command must survive. Using the model, it is possible to determine the desired survival probabilities for many combinations of the several variables: enemy-weapon CEP, enemy-weapon yield, enemy-weapon numbers, weapon-element and command-unit mixes, weapon-element or command-unit hardness, and weapon-element and command-unit dispersal.

2/ The range of dispersion distances within which this model will prove useful is from about 0.2 LR (lethal radii) to <2.0 LR. Sites dispersed less than about 0.2 LR can usually be considered as a single (point) target. The vulnerability of sites dispersed 2.0 LR or more can be determined through the use of the equations of the model and the probability tables of The RAND Corporation Report R-234, Offset Circle Probabilities. (1)

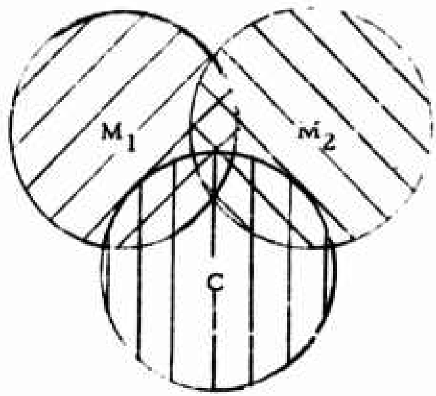
Conceptually, the method employed is simple. Diagrams of the weapon-complex layouts are constructed in the two-dimension-event-space defined by the probability-density function of enemy-weapon drop locations about an aiming point. The exhaustive set of mutually-exclusive states which the system can occupy are then delineated, e.g., left-hand missile and the command survive, no missile survives, etc. Next, the state-to-state transition probabilities are defined in terms of the simple events of the diagram previously constructed. These expressions are combined into a matrix of transition probabilities, which defines a Markov chain, and raised to the powers corresponding to the number of enemy weapons being considered. The resultant matrices delineate the probabilities of the mutually-exclusive ways in which each composite event can occur, where a composite event is a single state or combination of states, e.g., two missiles of three surviving an attack of three enemy weapons. To evaluate the composite event probabilities, the weapon-complex diagram is superimposed on a probability grid, and the probability of occurrence of each simple event is obtained by counting the included cells. These probabilities are then combined into the complex-event expressions.

2.0 TWO MISSILE SITES--ONE COMMAND SITE

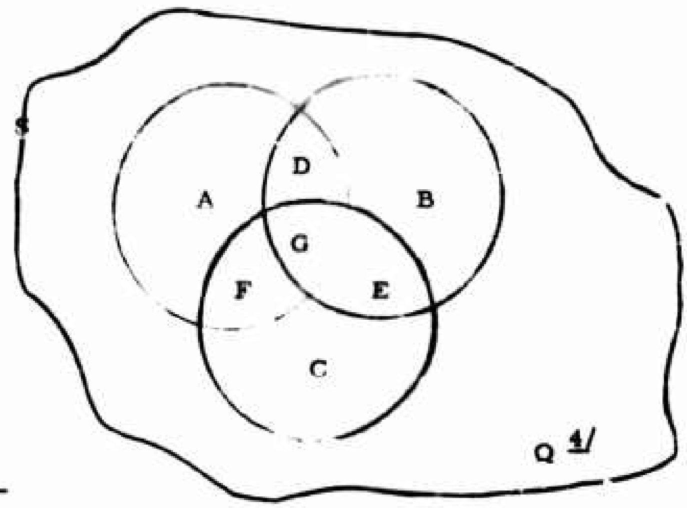
Consider a weapon complex composed of two missile sites (M_1 and M_2) and one command site (C). In order for a missile to be launched after an enemy attack, both it and the command site must survive the attack. All sites are hardened and capable of withstanding the same blast effects.[3/] (In this case the primary vulnerability design-factor is assumed to be overpressure.) For an enemy weapon of given yield, a lethal area can be defined about each site, within which an enemy weapon will be capable of covering the site with a destroying level of overpressure. If each of the sites is separated from the others by less than two lethal radii their

3/ This is not a constraint upon the model. Sites of unequal vulnerability can be handled with minor perturbations of the vulnerability diagrams.

vulnerable areas are as follows:



Let the missile-site and command-site vulnerable zones be labeled as follows, with each lettered sub-set being the smallest region enclosing the letter: S is the two-dimension event space corresponding geometrically to the area within which all bombs (or missiles) aimed at the arbitrary DGZ will fall. And, for example, D is that sub-set of the total event space with whose points is associated the occurrence of the simple event: destruction of both missile sites.



4/ Q is that sub-set associated with weapon complex invulnerability.

Before proceeding, several definitions are needed. First, the probabilities of occurrence of the simple events associated with each of the sub-sets, or vulnerable zones, are defined as:

$$a \equiv P(A)$$
$$b \equiv P(B)$$
$$\vdots$$
$$q \equiv P(Q),$$

where

$$a + b + c + d + e + f + g + q = 1.$$

Next, the several states possible for this complex are defined as follows.

$2 \equiv$ both weapons operative

$M_1 \equiv$ only the left weapon operative

$M_2 \equiv$ only the right weapon operative

$0 \equiv$ no weapons operative.

The state-to-state single-step transition probabilities can now be defined in terms of the simple-event probabilities determined by the vulnerable zones. By inspection:

$$p_{22} = q$$

i.e., the probability of going from state 2 (both weapons operative) to state 2 (both weapons operative) is equal to q, the probability that the enemy weapon falls outside the vulnerable zones of either element. Similarly,

$$p_{2M_1} = b$$
$$p_{2M_2} = a$$
$$\vdots$$
$$p_{M_2M_1} = 0$$
$$p_{00} = 1$$

These transition probabilities are constant from trial to trial. Therefore, according to Reference 2 they can be combined into a transition matrix which defines a Markov chain with constant transition probabilities. The generalization of this process to time dependent transition matrices is discussed in section 6.0 of this example.

$$P = \begin{vmatrix} q & b & a & 1 - (a + b + q) \\ 0 & b + q & 0 & 1 - (b + q) \\ 0 & 0 & a + q & 1 - (a + q) \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

where the first row and first column correspond to transitions concerning both missiles, the second row and second column refer to the left-hand missile, the third row and column refer to the right-hand missile and the last row and column refer to no operative missiles. (For example, the probability of going from state M_2 to state M_2 in one step is equal to $a + q$, the probability of an enemy weapon falling on either area A or Q. Obviously, the zero operative state is an absorbing barrier.

For the probabilities associated with each state after n enemy weapons have been independently delivered, the matrix need be raised to the nth power and the desired terms read from the first-row entries, since the system always begins with both missiles operative. Because of the large number of zeros, this row-into-column multiplication process is not difficult.

The results of 1, 2, and 3 attacking weapons will be given. These are the working equations of the model which define the probability of occurrence of each composite-event.

One attacking weapon: 5/

5/ In general,

$P^{(\gamma)}_{\delta/\beta}$ = probability of occurrence of the composite event-command and exactly δ weapon elements of a total β in the complex survive an attack of γ enemy weapons.

$$p_{0/2}^{(1)} = 1 - (a + b + q)$$

$$p_{1/2}^{(1)} = a + b$$

$$p_{2/2}^{(1)} = q$$

As can be seen, the composite-event--one missile operative--can occur in two mutually exclusive ways: either the left or right missile can survive. Consequently, $p_{1/2}^{(1)}$ is the sum of the two probabilities corresponding to the two included states. This same technique is used for the other composite events.

Two attacking weapons:

$$p_{0/2}^{(2)} = 1 - 2q(a + b) - a^2 - b^2 = q^2$$

$$p_{1/2}^{(2)} = a^2 + b^2 + 2q(a + b)$$

$$p_{2/2}^{(2)} = q^2$$

Three attacking weapons:

$$p_{0/2}^{(3)} = 1 - p_{1/2}^{(3)} - p_{2/2}^{(3)}$$

$$p_{1/2}^{(3)} = a(a + q)^2 + aq(a + q) + aq^2 + b(b + q)^2 + bq(b + q) + bq^2$$

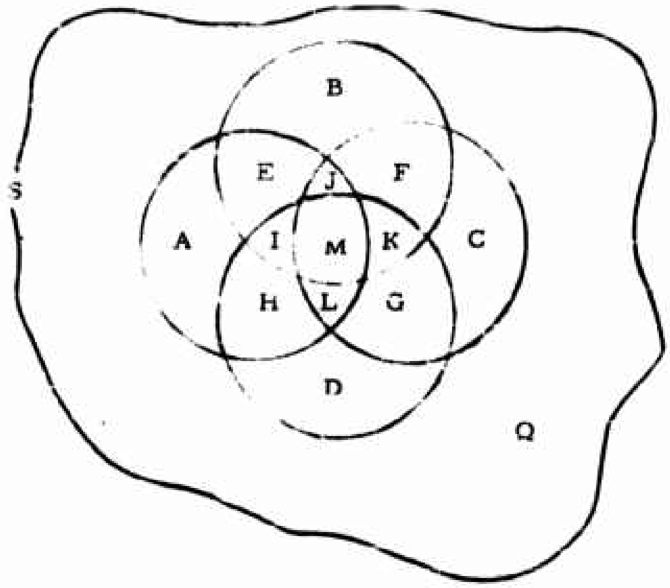
$$p_{2/2}^{(3)} = q^3$$

3.0 THREE MISSILE SITES--ONE COMMAND SITE

In order to examine a few more details of the model, consider an event-space for a left-hand missile (M_1), a middle missile (M_2), a right-hand missile (M_3) and a command unit at the bottom labeled as follows. (As can be seen, this layout is not a general case, but is specialized to the extent that the missiles are forward of the command.)

The several states which the system can occupy are:

3 = all three weapons operative



M_1M_2 ≡ left and middle weapons operative

M_1M_3 ≡ left and right weapons operative

M_2M_3 ≡ middle and right weapons operative

M_1 ≡ only the left weapon operative

M_2 ≡ only the middle weapon operative

M_3 ≡ only the right weapon operative

0 ≡ no weapons operative.

The state-to-state single-step transition probabilities can be obtained by inspection from the diagram. By combining these probabilities, the matrix of transition probabilities is defined. The row-and-column order is as above. And, for example, the probability of going from state M_1M_3 to state M_3, in one step, is equal to a + e, the probability of an enemy weapon falling on either area A or E.

The one-attacking-weapon survival probability equations for the several composite events are read directly from the first row of the matrix.

$$
P = \begin{vmatrix}
q & c & b & a & f & 0 & e & 1 - (a + b + c + e + f + q) \\
0 & c + q & 0 & 0 & b + f & a & 0 & 1 - (a + b + c + f + q) \\
0 & 0 & b + q & 0 & c + f & 0 & a + e & 1 - (a + b + c + e + f + q) \\
0 & 0 & 0 & a + q & 0 & c & b + e & 1 - (a + b + c + e + q) \\
0 & 0 & 0 & 0 & (b + c + f + q) & 0 & 0 & 1 - (b + c + f + q) \\
0 & 0 & 0 & 0 & 0 & (a + c + q) & 0 & 1 - (a + c + q) \\
0 & 0 & 0 & 0 & 0 & 0 & (a + b + e + q) & 1 - (a + b + e + q) \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{vmatrix}
$$

One attacking weapon:

$$p_{0/3}^{(1)} = 1 - (a + b + c + e + f + q)$$

$$p_{1/3}^{(1)} = e + f$$

$$p_{2/3}^{(1)} = a + b + c$$

$$p_{3/3}^{(1)} = q$$

As can be seen, $p_{1/3}^{(1)}$ is the sum of the probabilities of the three mutually exclusive ways for only one weapon to survive. Similarly, $p_{2/3}^{(1)}$ is the sum of the probabilities of the three mutually exclusive ways for two weapons to survive.

The P^2 (two attacking weapons) matrix, obtained through row-into-column multiplication of the P matrix, yields the two-attacking-weapon equations.

Two attacking weapons:

$$p_{0/3}^{(2)} = 1 - p_{1/3}^{(2)} - p_{2/3}^{(2)} - p_{3/3}^{(2)}$$

$$p_{1/3}^{(2)} = 2ab + 2ac + 2ae + 2bc + 2be + 2bf + 2cf + e^2 + 2eq + f^2 + 2fq$$

$$p_{2/3}^{(2)} = a^2 + b^2 + c^2 + 2aq + 2bq + 2cq$$

$$p_{3/3}^{(2)} = q^2$$

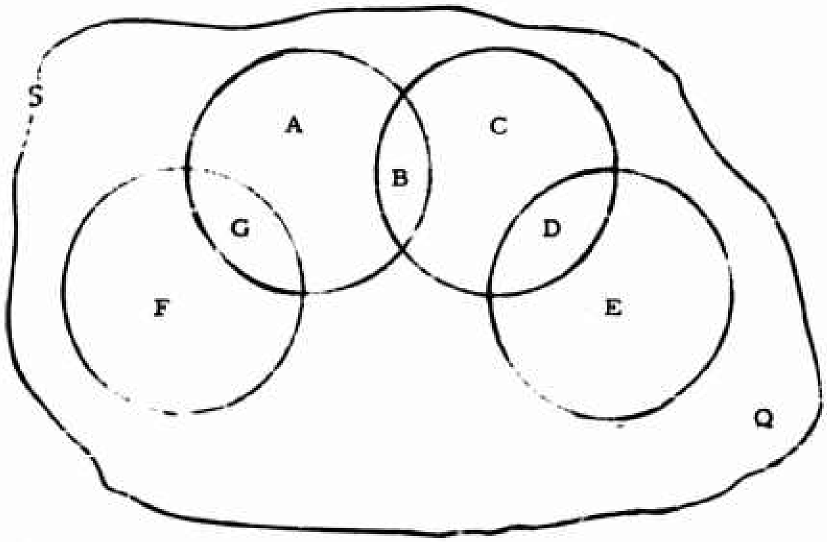
The following equations for three attacking weapons are obtained by row-into-column multiplication of the P and P^2 matrices. This shows the relative ease with which some rather awkward combinatorial expressions are obtained.

$$p_{0/3}^{(3)} = 1 - p_{1/3}^{(3)} - p_{2/3}^{(3)} - p_{3/3}^{(3)}$$

$$p^{(3)}_{1/3} = e^3 + f^3 + 3a^2(b + c + e) + 3b^2(a + c + e + f)$$
$$+ 3c^2(a + b + f) + 3f^2(b + c + q) + 3e^2(a + b + q)$$
$$+ 3q^2(e + f) + 6bcf + 6cfq + 6bfq + 6bcq + 6acq$$
$$+ 6beq + 6abq + 6abe + 6aeq$$
$$p^{(3)}_{2/3} = a(a + q)^2 + aq(a + q) + aq^2 + b(b + q)^2 + bq(b + q) = bq^2$$
$$+ c(c + q)^2 + cq(c + q) + cq^2$$
$$p^{(3)}_{3/3} = q^3$$

4.0 MORE THAN ONE COMMAND SITE

Preceding sections have considered complexes with only one command site; complexes with more than one command element can be handled as well. Consider, for example, a complex with two missile sites, M_1, M_2, forward of two command sites, C_1, C_2. Either command site can launch either missile.



The exhaustive set of possible states for this complex are:

$M_1M_2C_1C_2 \equiv$ both missiles and both commands operative

$M_1M_2C_1 \equiv$ both missiles and left command operative

$M_1M_2C_2 \equiv$ both missiles and right command operative

$M_1C_1C_2 \equiv$ left missile and both commands operative

$M_2C_1C_2 \equiv$ right missile and both commands operative

$M_1C_1 \equiv$ left missile and left command operative

$M_1C_2 \equiv$ left missile and right command operative

$M_2C_1 \equiv$ right missile and left command operative

$M_2C_2 \equiv$ right missile and right command operative

$0 \equiv$ no weapons operative

Proceeding as before, the transition matrix is as follows. This matrix can be employed to obtain the desired complex event expressions.

5.0 AN EXAMPLE

The development to this point has been general. No assumptions of probability-density function or aim point have been made; therefore, the transition-probability-matrices and resultant computing forms are valid for any choice of the foregoing. To provide an example of the computational details using a probability grid, a portion of the particular application for which the model was developed will be described.

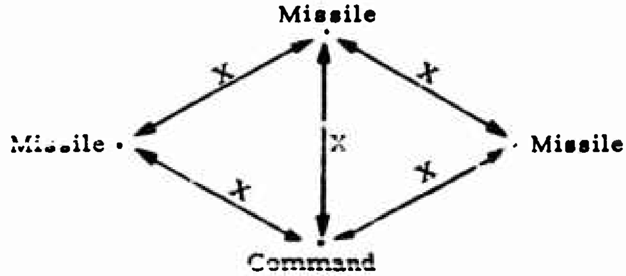
Consider a hardened missile-emplacement consisting of three missile sites and one command site. In order for either of the missiles to remain operative after an enemy attack, both the particular missile and the command must survive. Due to limitations of the guidance system, the missiles must be separated from the command by some minimum distance. In addition to the required separation, site dispersal is being considered to reduce vulnerability, but the cost of dispersal prohibits large separation. In order to determine the survival capability of the system as a function of

$$
P = \begin{vmatrix}
q & e & f & c & a & d & 0 & 0 & g & b \\
0 & q+e & 0 & 0 & 0 & c+d & 0 & a & 0 & b+f+g \\
0 & 0 & q+f & 0 & 0 & 0 & c & 0 & a+g & b+d+e \\
0 & 0 & 0 & q+c & 0 & d+e & f & 0 & 0 & a+b+g \\
0 & 0 & 0 & 0 & q+a & 0 & 0 & e & g+f & b+c+d \\
0 & 0 & 0 & 0 & 0 & q+c+d+e & 0 & 0 & 0 & a+b+f+g \\
0 & 0 & 0 & 0 & 0 & 0 & q+c+f & 0 & 0 & a+b+d+e+g \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & q+a+e & 0 & b+c+d+f+g \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q+a+f+g & b+c+d+e \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{vmatrix}
$$

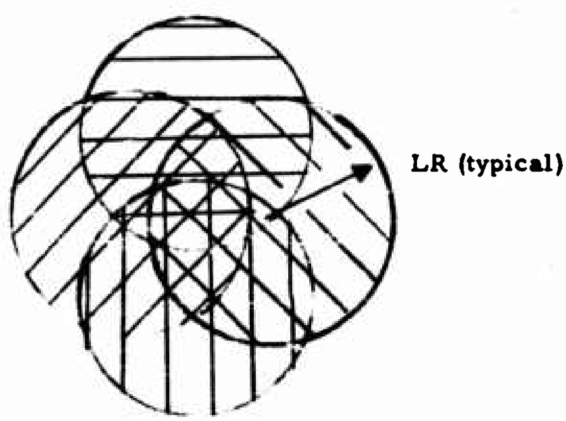
dollars spent on dispersal, it is required that the expected number of missiles remaining operative, after an attack by 1, 2, or 3 enemy weapons of specified size, be determined as a function of dispersal distance.

Each of the missile sites and the command are being designed to withstand the same overpressure. Then, for an enemy weapon of given yield, a lethal radius (LR) about each site can be defined, within which the site is vulnerable to the enemy weapon. It is desired to use this model to investigate the system vulnerability when each site is separated from the command by distances of less than two lethal radii. For reasons of control, missiles should be as close together as practicable; in the following sketch they are also separated by less than two lethal radii.

The complex whose vulnerability is to be estimated appears as follows:



For this layout, the vulnerable areas are:



where the dimension LR is the distance to which the design overpressure will extend from a given-yield weapon.

It is assumed that the delivery accuracy of the enemy weapon of interest can be described by the circular normal distribution where

$$\sigma_x = \sigma_y = \sigma .$$

For this distribution, CEP = 1.177σ.

Knowing the enemy delivery-accuracy, all dimensions can be transformed to multiples of the standard deviation, σ. After suitable transformations, the dispersion distances to be used in subsequent computations are: 0.23σ, 0.46σ, 0.69σ, 0.88σ, which correspond to 0.5 LR, 1.0 LR, 1.5 LR, and 1.9 LR, respectively. Survival probabilities will be computed for 1, 2, and 3 attacking weapons which will be assumed to be aimed at the command.

Because of the assumed enemy delivery-dispersion distribution, the two-inch standard deviation circular probability grid of H. H. Germond [3] is employed to determine the quantities of a, b, c, . . . of the probability equations. To do this, overlays of the several configurations to be studied are prepared using the scale of the probability grid. The overlays are then positioned over the grid with the assumed enemy aim-point (the command site) corresponding to the mean of the grid. This grid-overlay relationship is illustrated in Figure 1. The sites are dispersed 0.46σ (or 1.0 LR) from each other.

By counting the included elements, the quantities of the several sub-sets are seen to be:

a =	0.0370	h =	0.0175
b =	0.0225	i =	0.0165
c =	0.0365	j =	--
d =	0.0280	k =	0.0165
e =	0.0145	l =	--
f =	0.0145	m =	0.0070
g -	0.0175	q =	0.7720

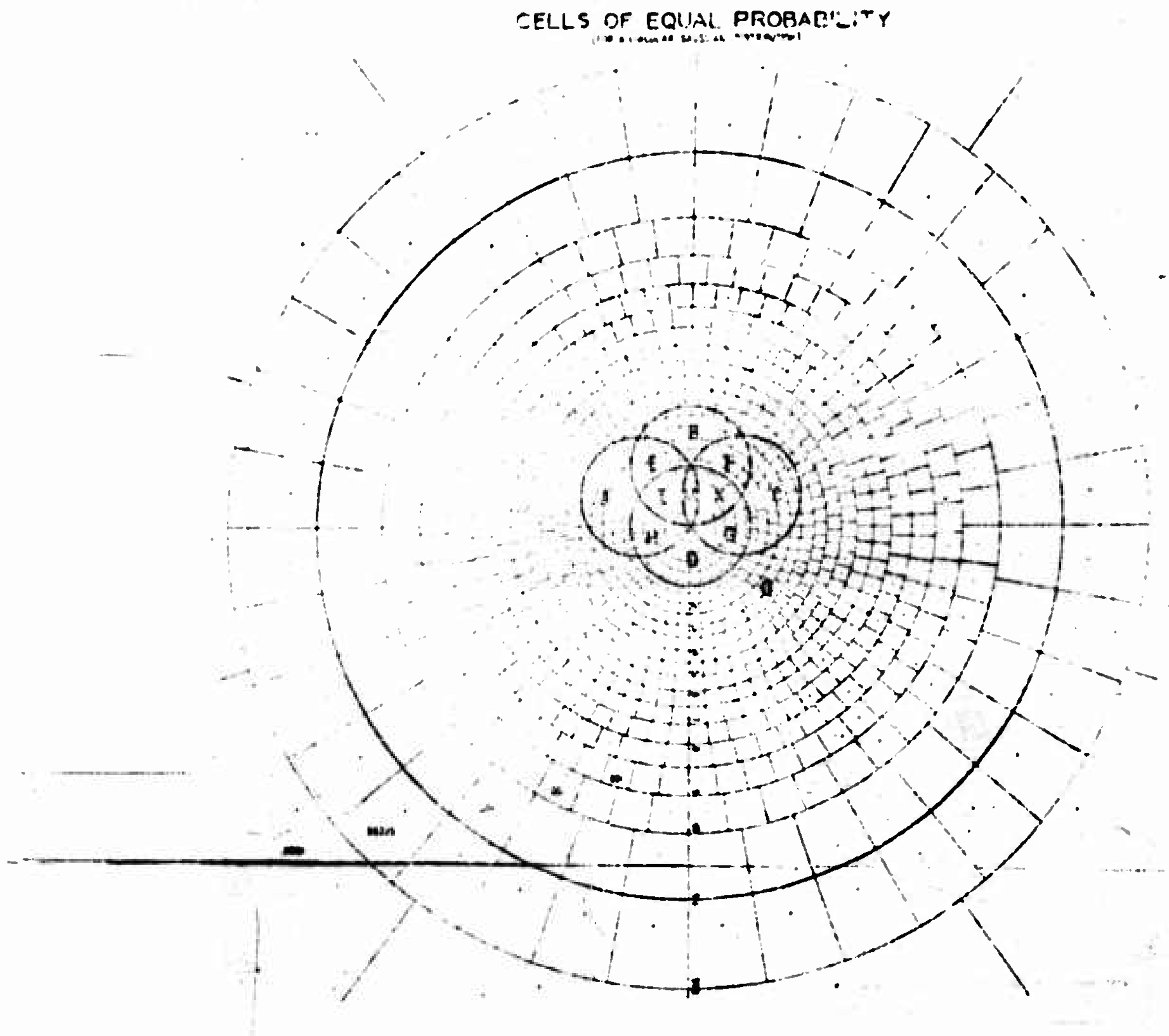


FIGURE 1

GRID-OVERLAY RELATIONSHIP SHOWING VULNERABLE REGIONS

These quantities, which are the probabilities associated with the occurrence of the several simple-events, are then combined into complex-event equations previously delineated, to determine the survival-probabilities. These are presented in Table I in the co x = 1.0 LR. The other survival probabilities shown in Table I we mined in the same manner.

The expected number of missiles surviving after an attack I calculated using the expression

$$R(N) = \sum_{N=1}^{3} N\, p_N$$

for each of the 12 combinations of dispersion distance and numb attacking weapons. The results are shown in Figure 2.

The data on Figure 2 for x = 0 and x = 2.0 are included fo of comparison. The x = 0 points were computed on the basis of a circular target; the x = 2.0 points were computed using the equa... this model and data from RAND Report R-234, Offset Circle Probabilities.(1)

6.0 GENERALIZATION OF METHOD

Previous sections have considered situations with only one type of enemy weapon delivered per complex and with only one aim point per complex. A complex with more than one command site, or a widely dispersed complex, may be best attacked by using two or more aim points. By using time-dependent variables, and by additional counting from probability grids, this model could be extended to accommodate both several aim points, and varying weapon characteristics, per complex. A general time-dependent Markov process would then be employed in place of the Markov chains with constant transition probabilities. The properties and limitations of the general Markov process are discussed in Reference 2. For this development, it is sufficient to note that whenever the conditions of the attack are changed from step to step--that is when more than one aim point is employed or when different enemy weapon characteristics are used for each weapon

PROBAB[illegible]IES

		Dispersion [illegible]stance (x) in [illegible] Radii			
		0.5	[illegible]	1.5	1.9
		0.104	[illegible]03	0.103	0.103
		[illegible]	[illegible]029	0.021	0.003
		[illegible]	[illegible]096	0.155	0.200
		[illegible]833	[illegible]772	0.721	0.694
[illegible] Attacking Weapons		[illegible]198	0.198	0.199	0.196
	$p_{1/3}^{(2)}$	[illegible]	0.055	0.049	0.032
	$p_{2/3}^{(2)}$	[illegible]071	0.151	0.232	0.291
	$p_{3/3}^{(2)}$	[illegible]695	0.596	0.520	0.481
Thre[illegible] Attacking Weapons	$p_{0/3}^{(3)}$	0.283	0.285	0.286	0.281
	$p_{1/3}^{(3)}$	0.[illegible]48	0.076	0.079	0.06[illegible]
	$p_{2/3}^{(3)}$	0.[illegible]90	0.179	0.260	0.31[illegible]
	$p_{3/3}^{(3)}$	0.[illegible]79	0.460	0.375	[illegible]

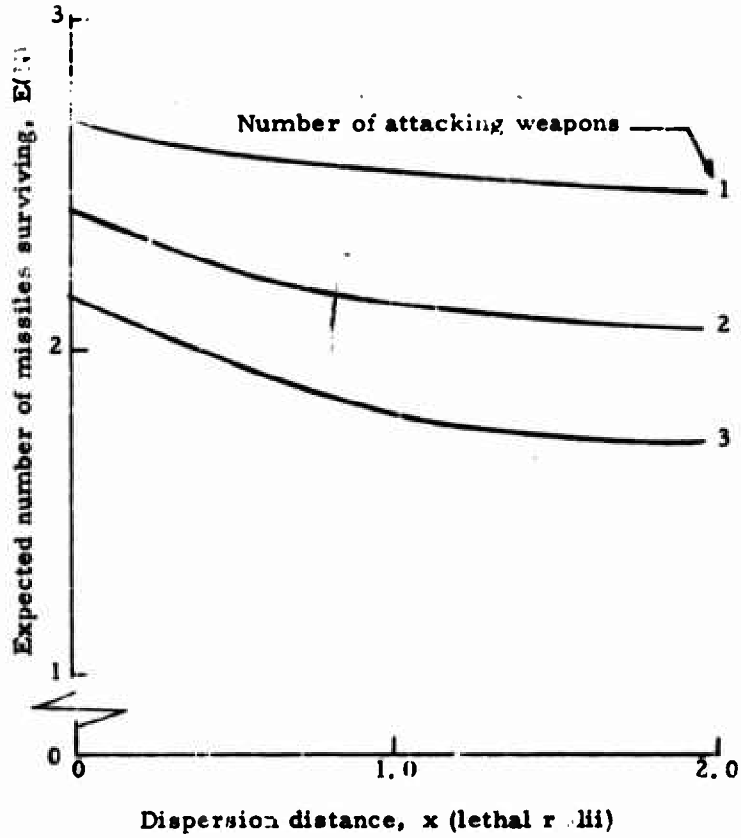


FIGURE 2

EXPECTED NUMBER OF MISSILE SITES SURVIVING AN ATTACK OF 1, 2, or 3 ENEMY WEAPONS (DGZ AT BLOCKHOUSE) delivered--the techniques described and the form of the transition matrix would remain unchanged. However, the quantities a, b, c, ... would become the time-dependent quantities, a_i, b_i, c_i, . . . the value of which would be dependent upon the step, i, in the process.[6/] This will lengthen the resultant probability expressions and will require a grid overlay exercise, peculiar to each step, to evaluate the probability expressions.

6/ In general, if a_o is the initial state probability vector, which by previous implication is given by

$$a_o = (1, 0, 0, \ldots, 0)$$

then for the Markov chain with constant transition probabilities, the state probability vector after n steps is given by

$$a_n = a_o P^n$$

where P is the matrix of constant transition probabilities. For the Markov process with time-dependent transition probabilities a_n is given by

$$a_n = a_o(P_1 P_2 P_3 \ldots P_n)$$

where P_i, (i = 1 , 2, 3, . . . n) is the matrix of transition probabilities for step i.

ES

E F]

'fset Circle Probabilities, The
March 14, 1952.

bility Theory and its
on New York, 1950.

T RAND
RM-309, Au 948

REFERENCES

1. Hitch, C., "An Appreciation of System Analysis," Journal of the Operations Research Society of America, November 1955.

2. Quade, E. S., "Military Systems Analysis," RM-3452-PR, The RAND Corp., January 1963.

3. Joyce, C. C., "Models: A Method for Systems Planning," Technical Documentary Report No. ESD-TDR-62-311, DDC No. AD 290539, The MITRE Corp., July 1962.

4. Leitman, George, Editor, Optimization Techniques. Academic Press, 1962.

5. Bellman, Richard, Editor, "Mathematical Optimization Techniques," R-396-PR, AD 409554, The RAND Corp., April 1963.

DISTRIBUTION L

Each WSEIAC Member	3
HQ USAF	
AFSPD	5
AFORQ	5
AFSLP	2
AFSME	2
AFOXP	2
AFRDC	2
Dep TIG USAF	5
AFRDP	2
SAC (OAOC)	25
ADC (ADMDC)	25
TAC (DMEMP/Maj J. Trent)	50
AFCS (CSSME/Col. Giettler)	10
CONAC (MME/Lt. Col. J. J. Harrington)	10
ATC (ATTAP)	15
AFLC (MCMCE)	75
AU (Library)	[illegible]
Air Force Academy	2
[illegible] (CE)	5
AFIT (SL)	100
HQ AF Systems Command	50
ASD (ASAPD/Mr. Overbe)	150
BSD (BSOC)	50
SSD (SSSIR)	50
ESD (ESTE)	50
FTD (TDE-1)	5
AMD (AMRO)	10
NRD	
SCGRS	1
SCGRP	1
NRD (Det #1)	
NRS	6
NRP	2

AFETR	[illegible]
AFWTR	[illegible]
ECMR	
RME	[illegible]
REQ	5
CCMR	
RCE	5
RCQ	[illegible]
WCMR	
RWE	5
RWQ	5
RTD (RTHC)	2
RADC (EMER)	30
SEG (SENPO)	40
RTD (Det #4)	5
AFWL (WLPP)	5
AFRPL (RPR)	5
AFSWC (SWTOP)	5
AEDC (AEL)	[illegible]
AFMDC (MDOPT)	5
AFFTC (FTTE)	5
APGC (PGVEC)	5
[illegible]D (DDR&E/Mr. Nucci)	50
Department of the Army OCRD (CRD/L) Wash., D. C.	5
Department of the Army OCRD (CRD/Q) Wash., D. C.	5
Office of Assistant Secretary of the Army Research and Development Attn: Mr. C. W. Woodside Wash., D. C.	3

Army Materiel Command (AMCRD-SA/Mr. L.A. Poepke) Wash., D.C.	25
Combat Developments Command (CDC-MR Col. R. J. Osborne) Ft. Belvoir, Va.	5
CNO (OP-701)	25
CNM (MAT-325)	200
BUSHIPS (Code 100-X)	25
BUWEPS (Code C)	50
ONR (Code 400)	20
Commandant, Marine Corps	20
NASA (Code KR)	30
National War College (Library)	1
Industrial College of the Armed Forces (Library)	1
Defense Management Senior SPO School (Col. Harris)	5
U. S. Army Management Engineering Training Agency (Rock Island Arsenal)	5
DDC	20